LOGIC

Jesper Carlström

2008

(revised English edition 2013)

Jesper Carlström
Matematiska institutionen
Stockholms universitet
106 91 Stockholm

http://www.math.su.se/~jesper/logikbok/
Translation from the Swedish original by Christian Espíndola and Håkon Robbestad Gylterud
Maintained by Peter LeFanu Lumsdaine; please report typos to p.l.lumsdaine@math.su.se
Last modified 2017-03-20

Typeset in LaTeX with the help of
Paul Taylor's package Prooftree

Contents

List of Figures

To the students

Many of the exercises are not collected at the end of the chapter, as usual, but inserted within the text itself. It is done precisely to point out when you should do them. They are there because, in my opinion, you need to check you have understood them before moving on. There are more exercises at the end of the chapter. Sometimes you will find a definition squeezed in amongst them. This is so because a logic student should practice also how to use definitions without having them explained first. Many of the exercises involves proving things. Logic is, after all, about proof.

Good luck!

Jesper Carlström
Stockholm
19 June 2007

To the teacher

This study material has been developed for the course Logic AN, 7,5 hp, at the Stockholm University Mathematics Department. The students at these courses have varied backgrounds, from mathematics, to philosophy and computer science. I have created this material with the intention of teaching the following in a basic course:

- Natural deduction in tree form.
- Mathematical properties such as soundness and completeness.
- Logic considered as a part of Mathematics.

The only book which was available and satisfied these criterions was van Dalen's *Logic and Structure*. This has been used for a long time, but was regarded by the students as too difficult. Teachers of the course thought it contained too many mistakes, and explained some parts in unnecessarily complicated ways, while it covered others too briefly. I have therefore decided to create new literature in the same spirit, but with a style that could be expected to suit a student better, so that more of them would complete the course.

When I got the assignment, I decided to improve the following aspects, compared to van Dalen's book:

- Set theory as the foundation of logic should be avoided since a natural application of logic is precisely set theory.
- The text should have a consistent terminology and a consistent notation to make the learning easier and more clear.
- Problems from exams should be included so that the students clearly see what is expected of them. (Many of the problems are taken from previous exams and are not constructed by me.)
- The text should be written in Swedish since it has been noticed that learning in ones mother language is much more efficient.
- The division of the material should be easy enough to understand so that it is clear what the student should do for every lecture, without having to give further reading instructions.

It is precisely because of this last item that I have adapted the content of the chapters, so that it is now suitable to take one chapter per lecture.

I will now comment on the material of certain chapters.

Chapters 1 and 2 deal with Boolean algebra. I consider it a natural introduction to symbolic logic for those who are used to thinking algebraically. The students that come from computer science usually recognize and treasure this part, which also contains references to computer science. Truth tables and normal forms are most easily described in an algebraic framework, and deciding whether a formula is a tautology, involves calculations in Boolean algebra. Last, but not least, Boolean algebra is an example of abstract algebra, and thereby gives the students experience with the notion of models before the subsequent chapters.

Chapter 3 treats inductively defined sets. The sets of formulas, terms, etc., which we later introduce and for which we give induction proofs, can be defined in this way. The advantage of a chapter about inductively defined sets is that one can discuss induction proofs and recursion in a natural way. It also avoids formulating a foundational system (set theory or type theory, for instance) for logic, and assumes instead a more structuralist approach: it is inductively defined sets what we need, while it is another (and in the context irrelevant) question in which framework we imagine that the theory about inductively defined sets should be formulated. The only sets we use in this course, which cannot be viewed as inductively defined sets, are the sets of equivalence relations, in Chapter 14. It is possible to disregard such sets by doing as Bishop did and letting equality be an equivalence relation rather than an identity relation, but I found such exposition too unfamiliar for the students.

In Chapter 6 the soundness of propositional logic is proved. Many books underestimate the importance of the soundness theorem, but I found it improper. Often, this is motivated by the fact that we have already argued for the derivation rules when they were introduced, so we know that they are sound. But these arguments are rarely solid, as they mainly serve as a sort of inspiration. It is, for example, far from obvious that the rules for undischarged assumptions are correct. In fact, one can look at the soundness theorem as a proof of this. Some students question the validity of the falsity-elimination, and are only convinced after seeing the proof of the soundness theorem.

Chapter 7 gives an introduction to normal derivations in propositional logic. The main purpose is to give the reader a tool for finding derivations in natural deduction in a methodical way, identifying which paths are dead ends. It is precisely for that reason that I have chosen to put the normalization proof in an appendix. Naturally, for a logician, it does not feel right to encourage students not to look at the proofs, but there are empirical indications that most students do not learn them since they are not expected to *normalize* when solving exercices, but only to search for normal derivations. The only reason why the normalization proof is in this course is the following: if something can be derived at all, it can be derived by a normal derivation. Students conceive the normalization proofs as difficult, but they often treasure the knowledge of how to search for normal derivations. This gives the possibility of answering precisely questions about which rules one “has to” use to derive a certain formula. To make the machinery of notions as easy as possible, I have chosen a definition of *normal* which is closely related to that of Seldin. It is useful for propositional logic, but less useful for predicate logic, since it is founded on Glivenko’s theorem. I have excluded normalization for predicate logic, since I think it suffices to have seen this for propositional logic, and because it is more complicated in the other case, with variable substitutions and everything.

Chapter 8 treats the completeness theorem. I have proceeded as in van Dalen’s book and chosen a proof which ressembles that of predicate logic as much a possible, with the intention of preparing them for this.

Chapter 10 introduces the semantics for predicate logic. Here I have chosen to set up a clear machinery for reevaluations to facilitate the understanding of how the truth value changes by substitution, as well as to make the soundness theorem easier. This is a big difference compared to van Dalen’s book, where substitutions take place in a completely informal way, and where the proof of the soundness theorem presents some difficulties.

Chapter 11 concerns how one “simplifies” formulas; that is, given a formula, how to find a new formula, which is simpler but has the same truth value as the original. This chapter also covers simplifying expressions which contain substitutions, through the use of reevaluations. The notion of *free for* is naturally introduced here.

In Chapter 12 the new rules which are needed for natural deduction are presented. I have chosen the less general rules, which do not allow changing variables when performing universal introduction and existential elimination, since I believe that the more general rules are too hard to grasp. In van Dalen's book *Logic and Structure* the simpler rules are used at first, but the more general rules are introduced towards the end of the book. However, the variable restrictions are formulated incorrectly, and the correct rules are more numerous and more difficult to check. The only place where the simple rules are a disadvantage in this book is in the proof of the model existence lemma, where one has to take a detour (a certain derivation becomes two steps longer before changing variables). If one wanted to perform normalization of predicate logic, it would be good to choose the more general rules, but since I have chosen to skip that topic one can restrict onself to the simpler rules.

I have included solutions to most of the exercises, except for the old exam problems, whose solutions are available at `http://www.math.su.se/`.

Jesper Carlström
Stockholm
Januari 2008, revised
October 2009

Thanks

Thanks to Clas Löfwall, who, during his period as prefect, made sure that a part of my teaching hours were reserved to writing this material. He also read an early version and provided many valuable remarks. Some of the problems included from old exams were constructed by him.

I thank Dag Prawitz for having given a critique of an early version of the section about normal derivations. It led me to rewrite the section completely. I hope that this made it considerably better.

Thanks also to Bengt Ulin, who helped me open my eyes to the pedagogical aspects of Boolean algebras.

Part I

Introduction

Chapter 1

Boolean algebra – Introduction

1.1 Boole's idea

Modern symbolic logic can be said to have started with an observation due to George Boole (1815–64). He noted that ordinary algebra can be used to formulate and solve logical problems. Think for example of t as "the tall", s as "the short", b as "the brown-haired".Then tb is interpreted as "the tall, brown-haired" and $t + s$ as "the tall and the short". One can formulate the principle that no one is both tall and short through the rule $ts = 0$ and then avoid problems by simplifying complicated expressions algebraically. For example, an unnecessarily complicated description of a collection as "the tall, which are brown-haired, but not the brown-haired short" which is symbolically represented by $t(b - bs)$ admits an algebraic simplication:

$$t(b - bs) = tb - tbs = tb - (ts)b = tb - 0b = tb \tag{1.1.1}$$

which shows that you can more easily call this group of people "tall and brown-haired".

In ring theory, in algebra, a ring is said to be *Boolean* if its elements satisfy the rule $a^2 = a$.

Boole also introduced the computation rule $aa = a$, or, to put it in another way, $a^2 = a$. This says, for example, that the brown-haired brown-haired people can be more easily described as brown-haired, and the short short people can be more easily described as short. In other words, it does not matter if you repeat a property several times. This computation rule can also be used to conclude that "those from the tall and the short which are short", can be more easily described as short:

$$(t + s)s = ts + ss = 0 + s = s\,. \tag{1.1.2}$$

Counting as Boole did has its problems, though. The following calculation fully complies with the usual computation rules, while also uses the rule $a^2 = a$:

$$2x = (2x)^2 = 4x^2 = 4x\,. \tag{1.1.3}$$

If one takes $2x$ from both sides to get $0 = 2x$, we conclude that the rule

$$a + a = 0 \tag{1.1.4}$$

must also hold. This says, if one applies it to s, that there are no people which are "short and short". Is this reasonable? Boole thought so. He simply did not allow adding twice the same expresion; in fact, he only allowed addition of two properties which are mutually exclusive. A problem with this idea is that one rarely knows whether two things can be added when you compute. A modern view is to accept addition of any two properties, but interpreting + as the *exclusive or*, which in computer language is often abreviated as XOR: if a and b are any two properties, then the property $a + b$ means of something that it has either property a or property b *but not both*. With this interpretation, it is quite reasonable to have the rule $a + a = 0$. One should just not read it as Boole did; instead of reading the expression $b + s$ as "the brown-haired and the short" one should read it as "either the brown-haired or the short, but not both". Since this version of "or" excludes the case of having both properties it is know as the *exclusive or.*

Unfortunately, sometimes the exclusive or is not very useful. If the data system of a ticket selling machine has indicated that discount should be offered to retired people or to students, we would still want those retired people that are students to receive the discount. If this is the interpretation of the word "or", it is called the *inclusive or* and often denoted with the symbol $\vee$ instead of $+$. Even in mathematics the inclusive or is preferable. People say things such as "if $a + b > 0$, then $a > 0$ or $b > 0$" but they do *not* exclude the case where both numbers can be positive.

The symbol $\vee$ comes from the first character of the latin *vel*, which means "or".

For the inclusive or we do have the rule $a \vee a = a$, just like $aa = a$. There is, therefore, some sort of similarity, more particularly a *duality*. To highlight this property one usually writes $a \wedge a$ instead of aa. One reads $\vee$ as "or" and $\wedge$ as "and". Hence, one reads $t \wedge b$ as "the tall and the brown-haired", while $t \vee b$ means "the tall or the brown-haired", but assuming one also includes the people that are both tall and brown-haired. As we saw earlier, we have automatically the rule $a + a = 0$ with the usual computations. We cannot, therefore, expect to count as with $+$ when we use $\vee$. You can almost always think of $\vee$ as $+$, but not really always. The problem is that you have no substraction; more precisely, there is not always a solution to the equation $a \vee x = 0$. Instead of substraction, one has the *complement*: one can write $\neg a$ for the property "not a". Hence, one writes, $\neg t$ for those who are not tall, and $\neg b$ for those who have are not brown-haired. The foundational rules for $\vee$, $\wedge$ and $\neg$ are not different from the usual properties that hold. They are collected in Figure 1.1. They are called axioms for Boolean algebra, even though Boole himself did not study this algebra: the name is used to emphasize that it is developped in a Boolean spirit. The list of axioms is unnecessarily long, since it is enough to have (comm), (id), (distr) and (inv) to derive the rest of the rules. It is convenient, however, to see all of them written down explicitly.

"Duality" refers to a pair that relate each other as direct opposite. Ironically, this supposes that they are actually strongly linked. Indeed, a duality requires that the pair consists of conformationally similar concepts, which are opposite in another sense. For example, the concepts of most and least can be said to be dual, but one would hardly say that most and yellow are dual.

Learn the rules of Figure 1.1 by heart, so it becomes much easier to solve problems. Use the help of duality and the names of the rules when you memorize them.

$a \vee b = b \vee a$	$a \wedge b = b \wedge a$	(comm)
$(a \vee b) \vee c = a \vee (b \vee c)$	$(a \wedge b) \wedge c = a \wedge (b \wedge c)$	(ass)
$a \vee 0 = a$	$a \wedge 1 = a$	(id)
$a \vee 1 = 1$	$a \wedge 0 = 0$	(abs)
$a \vee a = a$	$a \wedge a = a$	(idemp)
$a \vee (b \wedge c) = (a \vee b) \wedge (a \vee c)$	$a \wedge (b \vee c) = (a \wedge b) \vee (a \wedge c)$	(distr)
$a \vee \neg a = 1$	$a \wedge \neg a = 0$	(inv)
$\neg(a \vee b) = \neg a \wedge \neg b$	$\neg(a \wedge b) = \neg a \vee \neg b$	(dM)

comm stands for *commutativity*
ass stands for *associativity*
id stands for *identity element*
abs stands for *absorption*
idemp stands for *idempotence*
distr stands for *distributivity*
inv stands for *inverse element*
dM stands for *de Morgan's rules*

Figure 1.1: Axioms for Boolean algebras

1.1.5 Exercise. Which of the axioms for boolean algebras (Figure 1.1) are valid in usual algebra if we interpret $\vee$, $\wedge$ as $+$, respectively $\cdot$, and $\neg a$ as $1-a$?

1.1.6 Exercise. Show that the axiom (idemp) is not really needed.

> *Hint.* Begin by writing $a \wedge a$ as $(a \vee 0) \wedge (a \vee 0)$, then use (id) and afterwards (distr).

1.1.7 Exercise. Show that $\neg\neg a = a$ for every a.

> *Hint.* Show that $\neg\neg a = \neg\neg a \vee a$ and that $a = a \vee \neg\neg a$. Start, for example, as follows: $\neg\neg a \overset{\text{(id)}}{=} \neg\neg a \vee 0 \overset{\text{(inv)}}{=} \ldots$.

1.2 Examples of Boolean algebras

1.2.1 Example (trivial Boolean algebra)**.** The simplest example of Boolean algebra is so simple that is called *trivial*. It simply lets 0 and 1 just be the name of the same element $*$ and let $* \vee * = * \wedge * = \neg * = *$. This algebra just consists of a single element! As you might expect, it is not particularly useful. But it can serve to understand that $0 \neq 1$ does not follow from the Boolean algebra computation rules. Indeed, if it did, it should be valid in all Boolean algebras, while in the trivial algebra $0 \neq 1$ is false.

1.2.2 Example (algebra with two elements)**.** The simplest non trivial boolean algebra is obtained by considering the set $\{0, 1\}$ and defining the operations $\vee$, $\wedge$ and $\neg$ through a few simple tables. Think of 0 as representing *false* and 1 as representing *true*. Then, it is reasonable to set up the following so called *truth table*.

a	b	$a \vee b$
0	0	0
0	1	1
1	0	1
1	1	1

a	b	$a \wedge b$
0	0	0
0	1	0
1	0	0
1	1	1

a	$\neg a$
0	1
1	0

(1.2.3)

This algebra is often called the *initial Boolean algebra*.

1.2.4 Exercise. Check that the axioms of Figure 1.1 are satisfied if one defines the operations in (1.2.3).

1.2.5 Exercise. Show that the table (1.2.3) can only be completed in one way if we want the axioms of Figure 1.1 to hold. More specifically, show that the columns under $a \vee b$, $a \wedge b$ and $\neg a$ are fully determined by these axioms.

> *Hint.* It suffices to consider axioms (id), (abs), (inv) to show that the table is uniquely determined.

Despite its simplicity – or perhaps because of it – the two-elements-algebra is very important. It has applications in digital technology, but will also be basic for everything we do in this course. Once we introduce the *semantics* for propositional logic and predicate logic, it will be this algebra the one we will use (chapters 4 and 10).

1.2.6 Example (algebra generated by subsets)**.** There is another important algebra which is closer to what Boole wanted to do from the beginning. Consider all the students in a classroom. We can draw, on the classroom floor, three overlapping circles b, t, s (Figure 1.2), a so called *Venn diagram*. We ask now all brown-haired people to stand in the circle b, all tall people in the circle t, all short people in the circle s. Those who are brown-haired and tall can stand in the area where b and t overlap; that is, where the circular disks (the interior of the circles) *intersect* each other. Those who do not consider themselves to be brown-haired, tall or short can stand outside all circles. Where circular disks t and s intersect each other there should be no one, since it is not reasonable that someone is both tall and short. It is said, therefore, that the intersection is empty and we denote $t \cap s = \emptyset$. Note that names here are

Venn diagrams are named after the mathematician John Venn (1834–1923). This is a little unfair since Leonhard Euler used them already in 1768.

That the intersection of t and s is empty means that no one is in the space that t and s have in common.

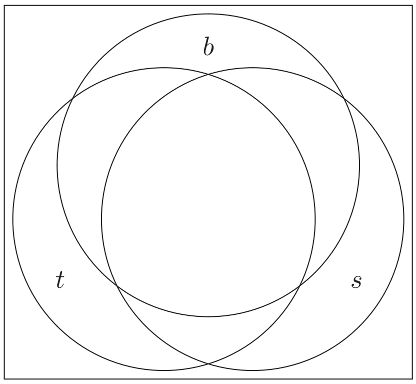

Figure 1.2: Venn diagram

somewhat different. We use $\cap$ instead of $\wedge$ and $\emptyset$ instead of 0. Dually, we can write $\cup$ instead of $\vee$ and often I instead of 1. One also often writes $\complement b$ or b^c instead of $\neg b$ and call this subset the *complement* of b – it is the subset of those that are not in b. The collection $b \cup t$ is called the *union of* b *and* t since you can think that joining the brown-haired and the the tall forms $b \cup t$.

If one disregards the fact that the names have changed a little, the axioms of Boolean algebras are satisfied (check some of them until you understand how they work). We have therefore a Boolean algebra with elements $\emptyset$ and I, but also b, t, s and all combinations between them, like $b \cup t$. One calls this the algebra *generated* from b, t, s.

1.2.7 Example (algebra of *all* subsets)**.** Given a set, we can consider the algebra of *all* its subsets. We interpret 0 as the empty set *emptyset*, 1 as the whole set I (the one that contains all elements of the original set), $\wedge$ as intersection $\cap$, $\vee$ as union $\cup$, $\neg$ as complement $\complement$. the axioms for boolean algebras are again satisfied, so this constitutes a new Boolean algebra. It consists of the *power set* (the set of all subsets) of the original set, together with the usual subset operations.

Let us now define precisely what a Boolean algebra is:

► **1.2.8 Definition.** A Boolean algebra is a set M, with constants $0 \in M$ and $1 \in M$, together with operations $\vee$ and $\wedge$ (binary) and $\neg$ (unary), such that the axioms of Figure 1.1 are satisfied.

1.3 Some properties of Boolean algebras

It is easy to check that $\neg 0 = 1$ and $\neg 1 = 0$ in the Boolean algebras we previously considered. Can we be sure that this holds in *every* Boolean algebra? For Boolean algebras in general, the only thing we know is that they fulfill the axioms. The answer is yes, which is shown by the fact that the following calculation is correct in all Boolean algebras:

$$\neg 0 \overset{\text{(id)}}{=} \neg 0 \vee 0 \overset{\text{(comm)}}{=} 0 \vee \neg 0 \overset{\text{(inv)}}{=} 1\,. \tag{1.3.1}$$

1.3.2 Exercise. Show that $\neg 1 = 0$ holds in all Boolean algebras.

Absorption rules:

$$a \vee (a \wedge b) = a$$
$$a \wedge (a \vee b) = a$$

1.3.3 Exercise. The following so called absorption rules hold: $a \vee (a \wedge b) = a$ and $a \wedge (a \vee b) = a$. The first one can be proved as follows:

$$\begin{aligned} a \vee (a \wedge b) &\overset{\text{(id)}}{=} (a \wedge 1) \vee (a \wedge b) \overset{\text{(distr)}}{=} a \wedge (1 \vee b) \\ &\overset{\text{(comm)}}{=} a \wedge (b \vee 1) \overset{\text{(abs)}}{=} a \wedge 1 \overset{\text{(id)}}{=} a\,. \end{aligned} \tag{1.3.4}$$

Prove the other.

A very useful principle when one works with Boolean algebras (especially when solving equations, as in Section 2.1) is that if $a \vee b = 0$, then both $a = 0$ and $b = 0$. This is also easily checked in the Boolean algebras we have considered. To be sure that it holds in any Boolean algebra, we use that the following is true if $a \vee b = 0$:

$$\begin{aligned} a &\overset{\text{(id)}}{=} a \wedge 1 \overset{(1.3.1)}{=} a \wedge \neg 0 = a \wedge \neg(a \vee b) \overset{\text{(dM)}}{=} a \wedge (\neg a \wedge \neg b) \\ &\overset{\text{(ass)}}{=} (a \wedge \neg a) \wedge \neg b \overset{\text{(inv)}}{=} 0 \wedge \neg b \overset{\text{(comm)}}{=} \neg b \wedge 0 \overset{\text{(abs)}}{=} 0\,. \end{aligned} \tag{1.3.5}$$

1.3.6 Exercise. Prove that, similarly, if $a \vee b = 0$, then $b = 0$. More precisely: prove that it is true in *all* Boolean algebras. Prove also that if $a \wedge b = 1$, then both $a = 1$ and $b = 1$.

We collect these useful results in the following theorem.

1.3.7 Theorem. *In a Boolean algebra, if $a \vee b = 0$, then $a = 0$ and $b = 0$. Dually, if $a \wedge b = 1$, then $a = 1$ and $b = 1$.*

The following exercise shows that this theorem should be read carefully:

1.3.8 Exercise. Show that in the Boolean algebra with two elements, if $a \wedge b = 0$, then $a = 0$ or $b = 0$. Give also an example of a Boolean algebra where this principle does *not* hold.

It does *not* always hold that if $a \wedge b = 0$ then a or b are 0.

Perhaps you have already thought about the fact that in our examples there is a *ordering* between the elements. In the Boolean algebra which consists of only 0 and 1, it is natural to say that $0 \leq 1$, and in the algebra of all subsets we have the relation $\subseteq$ which says that each element in a certain subset is also an element in another subset. For example, every person who is "brown-haired and tall" is also "tall", so we have $b \cap t \subseteq t$. There is in fact such an ordering in any Boolean algebra, which can be simply defined it in the following way:

► **1.3.9 Definition.** In a Boolean algebra, $a \leq b$ means that $a \wedge b = a$.

Note that in the Boolean algebra of subsets, $a \leq b$ is true precisely when $a \subseteq b$ is true.

1.3.10 Exercise. Show that $0 \leq 1$ holds, according to the definition.

1.3.11 Exercise. Show that $a \wedge b \leq b$ holds for any pair of elements a, b in any Boolean algebra.

When expressing oneself as in the statement of Exercise 1.3.11 one does *not* mean that you can *choose* a Boolean algebra and show that it holds there, but rather that you prove that it should hold in *every* Boolean algebra. The idea is that if I choose a Boolean algebra and two elements a, b then *you* should be able to show that $a \wedge b \leq b$ holds in it.

1.3.12 Exercise. Show that $\leq$ is what a mathematician calls a *partial ordering*:

$$a \leq a \qquad \text{(reflexivity)}$$

$$\text{If } a \leq b \text{ and } b \leq c, \text{ then } a \leq c\,. \qquad \text{(transitivity)}$$

$$\text{If } a \leq b \text{ and } b \leq a, \text{ then } a = b\,. \qquad \text{(antisymmetry)}$$

1.3.13 Exercise. Show that $\vee$ gives the *least upper bound* in the following sense:

$$a \leq (a \vee b)$$

$$b \leq (a \vee b)$$

$$\text{If } a \leq c \text{ and } b \leq c, \text{ then } (a \vee b) \leq c.$$

Hint. Here the absorption rules (exercise 1.3.3) comes into use.

1.3.14 Exercise. Show that $\wedge$ gives the *greatest lower bound*. Start by defining precisely what this means, analogously to the previous exercise.

1.3.15 Exercise. An *atom* is an element which is minimum amongst the elements which are not 0. In plain language: an $a \neq 0$ such that if $c \leq a$ for some $c \neq 0$, then $c = a$. Give examples of some atoms in some Boolean algebras. Prove that they are indeed atoms.

Isn't it wonderful what a mathematician calls plain language?

1.3.16 Exercise. Prove that if $a \leq b$, then $(a \vee c) \leq (b \vee c)$ for all c.

1.3.17 Exercise. Prove that if $a \leq b$, then $\neg b \leq \neg a$.

1.4 Precedence rules

Since we have associative rules in Boolean algebras, we do not need to write all the parentheses. For example, one has:

$$(((a \wedge b) \wedge (c \wedge d)) \wedge e) \wedge f = a \wedge (((b \wedge c) \wedge d) \wedge (e \wedge f))\,, \tag{1.4.1}$$

so it is enough to write:

$$a \wedge b \wedge c \wedge d \wedge e \wedge f\,. \tag{1.4.2}$$

The same happens with $\vee$. To further diminish the number of parentheses, one usually lets $\wedge$ "bind stronger" than $\vee$, in the same way as $\cdot$ binds stronger than $+$:

$$a \vee b \wedge c \tag{1.4.3}$$

means $a \vee (b \wedge c)$. Finally, $\neg$ binds stronger than $\wedge$.

1.4.4 Exercise. Simplify the following expression using Boolean algebra

a) $x \vee y \wedge y \vee \neg x$

b) $x \wedge y \vee y \wedge \neg x$

c) $\neg(\neg(x \wedge y) \vee x) \vee y$

1.5 Normal forms

In usual algebra one seldom accepts having expressions such as:

$$(x+3)(x-x) + x + (x \cdot x + x)x + (x \cdot x \cdot 3 + 4x)(x + 3x)\,. \tag{1.5.1}$$

As a rule, it is rewritten into

$$13x^3 + 17x^2 + x\,. \tag{1.5.2}$$

This polynomial is in a kind of *normal form*. In Boolean algebras, normal forms are important as well. In some sense they are even more important than in usual algebra, as they can be used to solve equations to an even larger extent. In Boolean algebras one has two sorts of normal forms: *disjunctive* respectively *conjunctive* normal form. An expression in disjunctive normal form can look as follows:

$$(\neg x \wedge y \wedge z) \vee (y \wedge z) \vee x \tag{1.5.3}$$

and an expression in conjunctive normal form can look like this:

$$(\neg x \vee y \vee z) \wedge (y \vee z) \wedge x\,. \tag{1.5.4}$$

> Compare:
>
> or: disjunction
> and: conjunction
> plus: addition
> times: multiplication

> Variables are denoted by $x, y, z, \ldots$. when letters $a, b, c, \ldots$ are used, we assume *arbitrary elements* in the algebra. The variable x is in both disjunctive and conjunctive normal form. We cannot know if a is in disjunctive normal form, since we do not yet know how the element a is written.

▶ **1.5.5 Definition.** An expression is in disjunctive normal form if it is a finite disjunction of finite conjunctions of variables and/or negated variables. Every variable may appear at most once in each conjunction. The expressions 0 and 1 are said to be in disjunctive normal form, though there are no other expressions that contain them.

▶ **1.5.6 Definition.** An expression is in conjunctive normal form if it is a finite conjunction of finite disjunctions of variables and/or negated variables. Every variable may appear at most once in each disjunction. The expressions 0 and 1 are said to be in conjunctive normal form, though there are no other expressions that contain them.

Just as an empty sum is 0 and an empty product is 1, we say that an empty disjunction is 0 and an empty conjunction is 1.

Compare:

$$\sum_{n=1}^{0} a_n = 0$$
$$\prod_{n=1}^{0} a_n = 1$$
$$\bigvee_{n=1}^{0} a_n = 0$$
$$\bigwedge_{n=1}^{0} a_n = 1$$

1.5.7 Example. The following are all in disjunctive normal form:

1. $x \vee y \vee \neg z$,
2. $x \vee (y \wedge \neg z) \vee w$,
3. $x \wedge y$.

None of the following are in disjunctive normal form:

1. $(x \vee y) \wedge z$,
2. $(x \wedge \neg x) \vee y$,
3. $0 \vee x$.

1.5.8 Exercise. Which of the following expressions are in disjunctive normal form? Which are in conjunctive normal form?

1. 0
2. $(x \vee y) \wedge z$
3. $x \wedge y \wedge z$
4. x
5. $x \vee \neg x$
6. $(x \vee \neg x) \wedge y$
7. $x \vee 0$
8. $x \vee (y \wedge 1)$
9. $a \vee b$ (trick question)

Every Boolean expression can be "written in disjunctive normal form" (and even in conjunctive normal form, which is completely dual). That is to say, in every Boolean algebra one can construct an expression which is in disjunctive normal form and which is equal to the one we started with. This can be done in the following way:

1. Use Exercise 1.1.7 to rewrite $\neg\neg a$ as a.
2. Use (distr) to rewrite expressions of the form $a \wedge (b \vee c)$ as $(a \wedge b) \vee (a \wedge c)$. Expressions of the form $(a \vee b) \wedge c$ are handled by first applying (comm) to get $c \wedge (a \vee b)$ and afterwards continuing with (distr).
3. Use de Morgan's laws to rewrite $\neg(a \wedge b)$ as $\neg a \vee \neg b$ and $\neg(a \vee b)$ as $\neg a \wedge \neg b$.
4. Use (inv), (abs), (comm) and (ass) to rewrite conjunctions which contain one variable, both negated and non negated, as 0 (for example, one rewrites $x \wedge y \wedge \neg x$ as 0).
5. Use (idemp), (comm) and (ass) to rewrite several occurrences of one negated variable into one, and similarly for non negated variables (for example, one rewrites $x \wedge \neg y \wedge x \wedge \neg y$ as $x \wedge \neg y$).
6. Use (comm) and (abs) to rewrite $a \wedge 0$ and $0 \wedge a$ as 0, and similarly $a \vee 1$ and $1 \vee a$ as 1.
7. Use (comm) and (id) to rewrite $a \vee 0$, $0 \vee a$, $a \wedge 1$ and $1 \wedge a$ as a.

Repeat these steps until none of them can be applied any further. Then you will have something in disjunctive normal form. I practice, one does not write down every step. For example, one writes $\neg\neg a$ as a without further justification, but one should remember the reason (Exercise 1.1.7) to keep a clear conscience. In the same way, one can rewrite $\neg(a \wedge b \wedge c)$ as $\neg a \vee \neg b \vee \neg c$ without specifying all the steps.

1.5.9 Exercise. Write the following in disjunctive normal form:

a) $x \wedge (y \vee (z \wedge x))$

b) $x \wedge \neg(y \vee \neg z) \wedge \neg(\neg y \wedge \neg z)$,

c) $\neg y \wedge \neg z \wedge \neg(x \wedge \neg(y \vee \neg z))$

Note the compact form of the table: under every variable we have put the values on the current row; under each operation we have put the value resulting from performing that operation. The three columns on the left can be completely ommitted, but one has then to keep in mind that if a variable occurs several times on the same row, (as is here the case with$x$), it has to have the same value in each occurrence.

Another way to convert an expression into disjunctive normal form is to write down the truth table of the expression and read the disjuntive normal form from it. We illustrate this with an example.

1.5.10 Example. Write $x \wedge (y \vee (z \wedge x))$ in disjunctive normal form.

Solution. We construct a truth table:

x	y	z	x	$\wedge$	(	y	$\vee$	(	z	$\wedge$	x	)	)
0	0	0	0	**0**		0	0		0	0	0		
0	0	1	0	**0**		0	0		1	0	0		
0	1	0	0	**0**		1	1		0	0	0		
0	1	1	0	**0**		1	1		1	0	0		
1	0	0	1	**0**		0	0		0	0	1		
1	0	1	1	**1**		0	1		1	1	1		
1	1	0	1	**1**		1	1		0	0	1		
1	1	1	1	**1**		1	1		1	1	1		

(1.5.11)

From the truth table we can see that the expression is true in the last three columns. If we construct an expression for each of these rows, we can later put them together; first we create an expression with has the value 1 on the third last row, but 0 on every other. We are able to do so by choosing the expression $x \wedge \neg y \wedge z$: for this to have the value 1, we must have precisely $x = 1$, $y = 0$, $z = 1$. For the last two rows we choose $x \wedge y \wedge \neg z$ respectively $x \wedge y \wedge z$. Finally, we combine these expressions with disjunctions:

$$(x \wedge \neg y \wedge z) \vee (x \wedge y \wedge \neg z) \vee (x \wedge y \wedge z)\,. \tag{1.5.12}$$

It can be likely the case that you got a shorter expression after doing Exercise 1.5.9 a. There is nothing wrong if one gets a different answer: normal forms are not unique in Boolean algebras. □

The previous method has a theoretical problem: indeed, we see that we always get an expression in disjunctive normal form, but how do we know that it is always equal to the one we started with in every Boolean algebra? The tables shows that as long as we replace the variables of the expression by 0 or 1, we will get an equality. But in many Boolean algebras, there are lots of other elements as well. How can we know that we will get equality even if we replace instead the variables by these? Theorem 2.1.39 will prove that this in fact works: it is enough to check by inserting 0 or 1 to be sure that the expressions are equal for every other replacement. This is a quite surprising property of Boolean algebras.

1.6 Simple equations

We will now investigate how to solve equations where the right hand side is 0. In the next paragraph we will build further on this by finding out how to deal with equations where the right hand side is something else.

The right hand side 0, the left hand side only conjunctions and negations: different methods in different Boolean algebras.

1.6.1 Example. Solve the equation $x \wedge y \wedge \neg z = 0$.

Solution. The solutions in the two elements algebra are easy to find. There they are given by all possible combinations except $(x, y, z) = (1, 1, 0)$. One can see this by solving the corresponding equation where one interchanges the 0 and the 1. In that case, Theorem 1.3.7 implies that one need to have $x = 1$, $y = 1$, $\neg z = 1$.

In other Boolean algebras one cannot necessarily describe the solutions that neatly; we will content ourselves analyzing what the equation says about a colletion where x, y, z are interpreted as three subsets, $\wedge$ is interpreted as $\cap$, and so on. In that case, the equation clearly says that the intersection between x, y and $\neg z$ is empty. Any collection which has that property can therefore be seen as a solution to the equation. One can take Venn diagrams as in Figure 1.2 as a guide, and colour the area which is empty according to the equation. □

The example is typical: all equations where the right hand side is 0 and the left hand side is a conjunction of variables and negated variables can be handled in the same way. If the left hand side is something else, one can always write it in disjunctive normal form and later apply Theorem 1.3.7 to get a system of equations of the previous type. We will also illustrate this with an example.

The right hand side 0, and arbitrary left hand side: write the left hand side in disjunctive normal form and continue as in the above example.

1.6.2 Example. Solve the equation $x \wedge (y \vee (z \wedge x)) = 0$.

Solution. We start by writing the left hand side in disjunctive normal form, according to (1.5.12), for example, so that we get the equation:

$$(x \wedge \neg y \wedge z) \vee (x \wedge y \wedge \neg z) \vee (x \wedge y \wedge z) = 0\,. \tag{1.6.3}$$

According to Theorem 1.3.7 this has the same solutions as the system of equations:

$$\begin{cases} x \wedge \neg y \wedge z = 0 \\ x \wedge y \wedge \neg z = 0 \\ \;\; x \wedge y \wedge z = 0\,. \end{cases} \tag{1.6.4}$$

In the Boolean algebra with two elements one finds the solutions by marking in a table those elements which are common for each of the equations in the system (for example, one writes the eight possible rows and cross out the ones which are impossible according to the three equations). In a Venn diagram one can colour the three areas which are empty according to the three equations. Any collection which is empty in all the couloured areas solves the equations. □

That one cannot precisely describe the solutions makes the situation analogous to the case of indetermined equation system in linear algebra. Sometimes the solution set for a system of equations is a whole plane of points; one cannot then give a unique solution, but has instead to consider the equation as solved when one has given the plane in a suitable way. In Boolean algebra, the most suitable way is often to give a number of conjunctions which shall be 0.

1.6.5 Example. In a database in a pharmaceutical company one has stored information about gender and illness history. Let x be the women, y be the men, and z those people that have insomnia. In a certain search, one needs to pick the people that fulfill the query $(x \vee \neg y) \wedge ((z \wedge (\neg x \vee y)) \vee \neg((x \wedge \neg y) \vee z)$. An employee in the company complains and claims that no one will be picked out in this query. Is he correct?

Solution. He claims that the company database material solves the equation:

$$(x \vee \neg y) \wedge ((z \wedge (\neg x \vee y)) \vee \neg((x \wedge \neg y) \vee z) = 0\,.$$

Let us look at the solution of this. We first write the left hand side in disjunctive normal form: $(x \vee \neg y) \wedge ((z \wedge (\neg x \vee y)) \vee \neg((x \wedge \neg y) \vee z) = (x \vee \neg y) \wedge ((z \wedge (\neg x \vee y)) \vee ((\neg x \vee y) \wedge \neg z) = (x \vee \neg y) \wedge (\neg x \vee y) \wedge (z \vee \neg z) = (x \vee \neg y) \wedge (\neg x \vee y) = (x \wedge y) \vee (\neg y \wedge \neg x)$. The equation is equivalent to the system:

$$\begin{cases} x \wedge y = 0 \\ \neg x \wedge \neg y = 0\,. \end{cases}$$

This means that the skeptic is correct if and only if the database contains: a) no one who is both a woman and a man b) no one who is neither a woman nor a man. □

1.6.6 Exercise. Solve the equation $x \wedge \neg(y \vee \neg z) \wedge \neg(\neg y \wedge \neg z) = 0$.

1.6.7 Exercise. Solve the equation $\neg y \wedge \neg z \wedge \neg(x \wedge \neg(y \vee \neg z)) = 0$.

1.7 Summary

You have encountered Boolean algebra, which was historically the first approach to formal mathematical symbolic logic. This will be useful to you, partly as a foundation for logic, and partly as an example of an abstract algebraic theory. That there are numerous Boolean algebras that fulfill the same axioms is an example of the fact that a theory can have many models, which is something that we will use further on in the course. In the next chapter you will learn more about how to solve equations in Boolean algebras. The most important thing to remember in the sequel is the ability to algebraically simplify Boolean expressions to be able to calculate the so called *truth values*. The ability to set up and apply truth tables is valuable when one decides whether a so called *formula* is true in every possible so called *interpretation*, something you will do many times during this course.

Chapter 2

Boolean equations and implications

2.1 Equations, inequalities and equation systems

Previously we have seen how to solve equations where the right hand side is 0; even those systems of equations of that type can be solved in the same way. We will see now how the same method can be applied to solve inequalities. An example will be soon presented, but first we need a lemma to show how one can replace inequalities by equations that have precisely the same solutions.

2.1.1 Lemma. *The inequality $a \leq b$ is equivalent to the equation $a \wedge \neg b = 0$.*

Proof. Assume that $a \leq b$, which means that $a \wedge b = a$. Put $\wedge \neg b$ on both sides; then we get $a \wedge b \wedge \neg b = a \wedge \neg b$. The left hand side can now be simplified using (inv) and (begr) to 0.

Assume, on the other hand, that $a \wedge \neg b = 0$. Then we have:

$$a \wedge b = (a \wedge b) \vee 0 = (a \wedge b) \vee (a \wedge \neg b) = a \wedge (b \vee \neg b) = a \wedge 1 = a\,, \tag{2.1.2}$$

that is $a \leq b$. □

2.1.3 Example. Solve the inequality $x \wedge y \leq y \wedge z$.

Solution. We start by asserting that the inequality is equivalent to the equation $x \wedge y \wedge \neg(y \wedge z) = 0$ according to the lemma. Then we can proceed as before: write the left hand side in disjunctive normal form

Inequality: exchange them with equations where the right hand side is 0.

$$\begin{aligned} x \wedge y \wedge \neg(y \wedge z) &= x \wedge y \wedge (\neg y \vee \neg z) && (2.1.4)\\ &= x \wedge ((y \wedge \neg y) \vee (y \wedge \neg z)) && (2.1.5)\\ &= x \wedge (0 \vee (y \wedge \neg z)) && (2.1.6)\\ &= x \wedge y \wedge \neg z && (2.1.7) \end{aligned}$$

and conclude that the solution is given by the equation:

$$x \wedge y \wedge \neg z = 0\,. \tag{2.1.8}$$

□

With the help of inequalities we can now solve arbitrary equations. Equations of the form $a = b$ can be written, with the help of antisymmetry (Exercise 1.3.12), as a system of inequalities

Arbitrary equations: rewrite the system with two inequalities and proceed using the method for solving inequalities.

$$\begin{cases} a \leq b \\ b \leq a \end{cases} \tag{2.1.9}$$

which can then be solved as we did above.

2.1.10 Example. Solve the equation $x \wedge y = y \wedge z$.

Solution. We start by rewriting the equation into the system:

$$\begin{cases} x \wedge y \leq y \wedge z \\ y \wedge z \leq x \wedge y\,. \end{cases} \tag{2.1.11}$$

We have seen already in the previous example that the upper inequality has the same solutions as $x \wedge y \wedge \neg z = 0$. Similarly, we can show that the lower inequality has the same solutions as $\neg x \wedge y \wedge z = 0$. In the algebra of two elements, all combinations except $(1, 1, 0)$ and $(0, 1, 1)$ are solutions to the equation. In general, we cannot give a better answer than these two equations. □

Systems of equations: no more difficult than equations. One just get a system of equations bigger than the system we started with.

Now there are no difficulties in handling systems of equations and systems of inequalities; one just applies precisely the same methods. However, it is sometimes an advantage to apply the following method for solving an inequality rather than applying Lemma 2.1.1: given an inequality $LHS \leq RHS$, rewrite the LHS in dsjunctive normal form and the RHS in disjunctive normal form. Use now that $a \vee b \leq c$ is equivalent to the system:

$$\begin{cases} a \leq c \\ b \leq c \end{cases} \tag{2.1.12}$$

according to Exercise 1.3.13. Dually, $a \leq b \wedge c$ is equivalent to the system

$$\begin{cases} a \leq b \\ a \leq c\,. \end{cases} \tag{2.1.13}$$

In this way we can rewrite one large inequality into many small ones.

2.1.14 Example. Solve the inequality $(x \wedge y) \vee z \leq (\neg y \vee z) \wedge \neg w$.

Solution. Here we do not have to write in normal form because the left hand side is already written in disjunctive normal form and the right hand side is already given in conjunctive normal form. It follows immediately that the inequality is equivalent to the system:

$$\begin{cases} x \wedge y \leq \neg y \vee z \\ \quad\;\; z \leq \neg y \vee z \\ x \wedge y \leq \neg w \\ \quad\;\; z \leq \neg w\,. \end{cases} \tag{2.1.15}$$

The first inequality can be written in equational form: $x \wedge y \wedge \neg(\neg y \vee z) = 0$, which, when the left hand side is written in disjunctive normal form, becomes $x \wedge y \wedge \neg z = 0$.

The second inequality in the system is always true, since the right hand side is greater than the left hand side (see Exercise 1.3.13). Hence, we can ignore this one.

The third inequality in the system is equivalent to the equation $x \wedge y \wedge w = 0$.

The fourth inequality in the system is equivalent to the equation $z \wedge w = 0$.

The original inequality is thus equivalent to the system:

$$\begin{cases} x \wedge y \wedge \neg z = 0 \\ x \wedge y \wedge w = 0 \\ z \wedge w = 0\,. \end{cases} \tag{2.1.16}$$

We have already an answer in a good form. The complicated inequality has changed into three conditions which are considerably easier to check and to understand. Furthermore, one can get rid of one of them. Since variables x

and y occur in the first two equations, we can simplify a bit more. We rewrite the second equation as:

$$x \wedge y \wedge (z \vee \neg z) \wedge w = 0 \tag{2.1.17}$$

which, when we rewrite the left hand side in disjunctive normal form, gives the equation system:

$$\begin{cases} x \wedge y \wedge z \wedge w = 0 \\ x \wedge y \wedge \neg z \wedge w = 0 \end{cases} \tag{2.1.18}$$

Now we see that the upper equation follows from the last equation in (2.1.16), while the second one follows from the first equation in (2.1.16). Clearly, the second equation in (2.1.16) follows from the other two, so the latter one is enough. We can therefore answer that our original inequality is equivalent to the system:

$$\begin{cases} x \wedge y \wedge \neg z = 0 \\ z \wedge w = 0\,. \end{cases} \tag{2.1.19}$$

□

2.1.20 Example. Solve the inequality $(\neg y \vee z) \wedge \neg w \leq (x \wedge y) \vee z$.

Solution. We write the left hand side in disjunctive normal form and the right hand side in conjunctive normal form:

$$(\neg y \wedge \neg w) \vee (z \wedge \neg w) \leq (x \vee z) \wedge (y \vee z)\,. \tag{2.1.21}$$

Now we can rewrite the inequality as the following system:

$$\begin{cases} \neg y \wedge \neg w \leq x \vee z \\ z \wedge \neg w \leq x \vee z \\ \neg y \wedge \neg w \leq y \vee z \\ z \wedge \neg w \leq y \vee z\,. \end{cases} \tag{2.1.22}$$

The second inequality is always true, as $z \wedge \neg w \leq z \leq x \vee z$ (see Exercise 1.3.13, and dually for conjunction). Likewise for the fourth equation. The other two can be written in equational form:

$$\begin{cases} \neg y \wedge \neg w \wedge \neg(x \vee z) = 0 \\ \neg y \wedge \neg w \wedge \neg(y \vee z) = 0 \end{cases} \tag{2.1.23}$$

which, when the left hand side is written in disjunctive normal form, becomes:

$$\begin{cases} \neg y \wedge \neg w \wedge \neg x \wedge \neg z = 0 \\ \neg y \wedge \neg w \wedge \neg z = 0\,. \end{cases} \tag{2.1.24}$$

Here we see that the upper equation follows from the lower one, so only the lower one is relevant. The original inequality is thus equivalent to:

$$\neg y \wedge \neg w \wedge \neg z = 0\,. \tag{2.1.25}$$

□

2.1.26 Example (from the exam on 2007-08-17)**.** Solve the equation $(y \wedge x) \vee (x \wedge z) = x \wedge (x \vee z)$.

Solution. We first simplify the right hand side to x (absorption rule). The equation can now be written as a system of inequalities:

$$\begin{cases} (y \wedge x) \vee (x \wedge z) \leq x \\ x \leq (y \wedge x) \vee (x \wedge z)\,. \end{cases} \tag{2.1.27}$$

In the first inequality, the left hand side is in disjunctive normal form, so it can be rewritten as the system:

$$\begin{cases} y \wedge x \leq x \\ x \wedge z \leq x \end{cases} \tag{2.1.28}$$

which is solved by using the definition and instances of idempotence. We can therefore ignore this one. The original equation is thus equivalent to the inequality:

$$x \leq (y \wedge x) \vee (x \wedge z)\,. \tag{2.1.29}$$

We write the right hand side in conjunctive normal form:

$$x \leq x \wedge (y \vee z)\,. \tag{2.1.30}$$

This inequality is equivalent to the system:

$$\begin{cases} x \leq x \\ x \leq y \vee z\,. \end{cases} \tag{2.1.31}$$

The first of these equalities is always satisfied, so we can ignore it. The original equation is thus equivalent to the inequality:

$$x \leq y \vee z\,. \tag{2.1.32}$$

One cannot answer the question in a simpler way than this. Possibly, one prefers to write the inequality as an equation:

$$x \wedge \neg(y \vee z) = 0 \tag{2.1.33}$$

which can be simplified into

$$x \wedge \neg y \wedge \neg z = 0\,. \tag{2.1.34}$$

□

When you do the exercises below, you can try to solve the inequalities which arise both with the above method and through a direct application of Lemma 2.1.1.

2.1.35 Exercise. Solve the inequality $x \wedge y \leq z$.

2.1.36 Exercise. Solve the equation $x \wedge \neg(y \vee \neg z) = \neg y \wedge \neg z$.

2.1.37 Exercise. Solve the following system of equations and inequalities:

$$\begin{cases} x \wedge \neg(y \vee \neg z) = \neg y \wedge \neg z \\ x \wedge y \leq z \\ y \wedge z = 0 \end{cases}$$

2.1.38 Exercise (from the exam on 2007-01-10)**.** Solve the equation $x \wedge (y \vee z) = (y \vee z) \wedge (x \vee y)$

The methods we have presented in this section can also be used to prove the following useful theorem:

2.1.39 Theorem. *If an equation is satisfied when its variables are substituted by 0 and 1, it is also satisfied by all the elements of any Boolean algebra.*

Proof. Assume we have an equation which is satisfied when variables are substituted by 0 and 1. Apply the methods we have seen so far to write the equation as a system of equations where the right hand side is 0 and the left hand side is a conjunction of variables and negated variables. If all left hand sides are 0 (that is, we have empty conjunctions) then we are done, since the equation we started with is equivalent to $0 = 0$. Assume, therefore, that some left hand side contains a variable. Since every insertion of 0 and 1 makes this left hand side equal to 0, there must be some variable which occurs both negated and non negated in it, otherwise we could choose insertion of 0 and 1 for each variable so the left hand side is not 0, contradicting the fact that the equation is solved by all substitutions. But if a variable occurs both negated and non negated, the whole left hand side can be written as 0 by using (ass), (komm), (inv) and (id). To conclude, we have that only using the axioms of Boolean algebras we could show that the equation we started with is equivalent to $0 = 0$; that is, every insertion of elements of the Boolean algebras into the variables solves the equation □

The proof is quite compact; it is not important to learn it by heart, but the theorem is important in itself, since it shows that the methods of truth tables are useful when writing expressions in disjunctive normal form (Example 1.5.10).

2.2 Implication

Let us, as an introduction, consider a little bit informally a Boolean algebra of *conditions*. You can think about them as conditions for picking out entries in a database, but just as well as conditions for specifying a subset in mathematics: the condition *odd* gives, for instance, the odd numbers as a subset of the natural numbers. If a and b are two conditions, then the condition $a \wedge b$ is satisfied precisely when both a and b are satisfied. The condition $a \vee b$ is satisfied precisely if at least one of the conditions a and b are satisfied. The condition 0 is that which is never fulfilled, while the condition 1 is that which is always fulfilled. Two conditions are said to be equal if they are satisfied on the same set of things.

Now let a, b, c be three conditions and assume that the following has been observed:

$$\text{Everything which fulfills conditions } a \text{ and } b \text{ fulfills condition } c. \tag{2.2.1}$$

We then naturally draw the conclusion:

$$\text{Everything which fulfills the condition } a \text{ fulfills that if } b \text{ then } c. \tag{2.2.2}$$

Indeed, we know that if condition a is fulfilled, we then know that if b is fulfilled, according to (2.2.1), condition c will be fulfilled. Conversely, we can go from observation (2.2.2) to (2.2.1), since if a and b are fulfilled then a is fulfilled, and then, according to (2.2.2) that if b, then c; hence, since b is fulfilled it follows that c is fulfilled. We have therefore observed an equivalence between the principles (2.2.1) and (2.2.2).

In the language of Boolean algebra we can express (2.2.1) as $a \wedge b \leq c$, but (2.2.2) cannot be so easily expressed, since we do not have any symbols for *if... then...*. We will introduce further below such a symbol $\rightarrow$, and call the corresponding operation *implication*. We shall do this so that the equivalence between (2.2.1) and (2.2.2) can be expressed as

$$(a \wedge b) \leq c \iff a \leq (b \rightarrow c)\,. \tag{2.2.3}$$

Such a connection between $\wedge$ and $\rightarrow$ is in mathematics called a *Galois connection*. This kind of connections occurs in many places in mathematics.

We will now introduce an implication that fulfills (2.2.3), which in the algebra of conditions will work as a proper formal correspondence to *if... then...*, even though we will see that it has certain properties that one does not normally associate to *if... then...*. In other Boolean algebras we cannot expect that such an implication will correspond to the normal use of *if... then...* in a great extent: there are infinitely many Boolean algebras which are not related to conditions, but it was exactly the example of conditions what we have used to intuitively motivate implication. For example, in the two elements algebra, *if... then...* is a pretty far-fetched interpretation; what does “if 0 then 1” mean? We do not use this kind of sentences in our everyday language. The motivation we can give in *general* to introduce implication is that a Galois connection is certainly a good thing to introduce. That is shown by experience in all areas of mathematics. In different Boolean algebras, the interpretation of $\to$ will be different, but the Galois connection will always be there. In the algebra of conditions, the Galois connection captures exactly the important equivalence between (2.2.1) and (2.2.2). Further below we will prove that the operation $\to$ can always be defined in a way that one really gets the Galois connection with $\wedge$. For a start, you can investigate by yourself how it has to be in the two elements algebra.

2.2.4 Exercise. Investigate how $\to$ must work in the case of the two elements Boolean algebra by studying the case $a = 1$ in (2.2.3). Draw up a truth table for $\to$ as in (1.2.3).

We will now address the problem of how to introduce implication once and for all, by doing it simultaneously in every Boolean algebra. First we notice that the left hand side in (2.2.3) can be rewritten as an equation: $a \wedge b \wedge \neg c = 0$. This equation can be rewritten as $a \wedge \neg(\neg b \vee c) = 0$, which then can be expressed as the following inequality: $a \leq \neg b \vee c$. We can therefore express (2.2.3) equivalently as:

$$a \leq (\neg b \vee c) \iff a \leq (b \to c)\,. \tag{2.2.5}$$

This is naturally fulfilled if $(b \to c) = \neg b \vee c$, so one solution could be to simply define $(b \to c) \overset{\text{def}}{=} \neg b \vee c$. But perhaps there are other better ways? No, it is certainly not the case: we *must* have $(b \to c) = \neg b \vee c$ if (2.2.5) shall be valid for all choices of a. Indeed, if we let $a = (b \to c)$ in (2.2.5) we get $(b \to c) \leq (\neg b \vee c)$ and if we let $a = (\neg b \vee c)$ we get $(\neg b \vee c) \leq (b \to c)$. Since $\leq$ is shown in both directions, we get the equality. Because of this, we state the following definition:

▶ **2.2.6 Definition.** In a Boolean algebra, we define $a \to b$ as $\neg a \vee b$.

Precedence rules: $\to$ has lower priority than $\wedge$ and $\vee$, thus $x \wedge y \to z \vee w$ means $(x \wedge y) \to (z \vee w)$.

One usually reads $a \to b$ as "if a, then b", even when it does not have any immediate intuitive meaning. It may feel strange to say "if 0, then 1", but in Boolean algebras one often uses that expression. Remember that only in some special cases we have made attempts to capture something intuitive using this. It is the Galois connection (2.2.3) the property of mathematical importance, and the one which one looks for when introducing implication. The interpretation of *if... then...* is less important. If one thinks that this is a point of view excessively formal, since Boolean algebra is about truth values, one can think of $a \to b$ as "b is at least as true as a".

$a \to b$ can be thought of as "b is at least as true as a".

$a \to b = 1$ if and only if $a \leq b$.

2.2.7 Example. If 0 and 1 are truth values, where 0 represents false and 1 represents true, then 1 is at least as true as 0, and thus $0 \to 1$ is true.

2.2.8 Example. If b, t, s stands for brown-haired, tall, respectively short, then a person has the property $b \to t$ if the fact that he is tall is at least as true as the fact that he has brown hair. In other words, all blondes belong to $b \to t$, no matter their height, as it is false that they are brown-haired.

2.2.9 Exercise. Do you think it feels correct to say about a short blonde that if she is brown-haired then she is tall? Only you have the correct answer to this exercise.

2.2.10 Example. "If he is the king then I am Donald Duck" is something one could say. With this phrase, one might just mean that it is at least as true that I am Donald Duck as he is the king.

Let us now see how one can use implication in the most intuitive interpretations: when dealing with conditions. Let us take a concrete example.

2.2.11 Example. In a database of numbers from an experiment one wants to pick out those numbers a that satisfy the conditions:

- a is divisible by 3,
- if a is divisible by 2, then it is divisible by 4,

We let x be divisibility by 3, y be divisibility by 2 and z be divisibility by 4. The condition we should use can therefore be given as $x \wedge (y \to z)$. The content of the database is as follows: $1, 3, 6, 4, 12$. We shall investigate each and every of these numbers regarding the condition $x \wedge (y \to z)$. The numbers must satisfy the condition x and the condition $y \to z$ to be picked out. When we investigate if the number 1 satisfies the conditions, we discover immediately that condition x is not satisfied, so this number is not picked out. The next number is 3. Here the condition x is fulfilled, so it has thus far survived" our criteria. The next condition is $y \to z$; that is, *if* it is divisible by 2 *then* it must be divisible by 4. But we do not have divisibility by 2, so we are allowed to say that this condition do not give us any problems either: the number 3 has complied with our criteria and is then picked out. Next number is 6. Here condition x is satisfied, as well as y, but z is not satisfied. Hence, the criteria that *if* y *then* z fails. Therefore the number 6 is *not* picked out. Next number, 4, fails to satisfy the first condition, x, and hence it is not included. However, when we get to the number 12 we find that all the conditios x, y, z are satisfied, and hence, also the condition $y \to z$ is fulfilled. Therefore, the number 12 is picked out. The numbers we have picked out are then $\{3, 12\}$. Note that in this reasoning, one intuitively thinks about the condition $y \to z$ as only being *relevant* when y is fulfilled. When y is not fulfilled, one thinks that one can *skip* this condition. In a process when one filters elements that do not satisfy certain conditions, *skipping* is in practice the same as saying that the condition is *fulfilled*. Instead of saying: "if y is not fulfilled one skips the condition $y \to z$" one can say "if y is not fulfilled, one regards $y \to z$ to be satisfied". A moment of thought shows that the condition $y \to z$ will thus always be fulfilled whenever $\neg y \vee z$.

2.2.12 Exercise. Check that $(a \to 0) = \neg a$ holds in any Boolean algebra.

2.2.13 Exercise (from the exam on 2005-08-23)**.** Write the following expressions in disjunctive normal form: $\neg(\neg(x \wedge z) \to x) \to ((z \vee y) \wedge y)$.

▶ **2.2.14 Definition.** By $a \leftrightarrow b$ we mean that $(a \to b) \wedge (b \to a)$.

$a \leftrightarrow b$ can be thought of as "a and b are equally true". One calls $\leftrightarrow$ *equivalence*. It has as low precedence as $\to$.

$a \leftrightarrow b = 1$ if and only if $a = b$.

2.2.15 Exercise. Make a truth table for $\leftrightarrow$. One column for a, one column for b and one column for $a \leftrightarrow b$.

2.2.16 Exercise. Make a truth table for the following expression:

$$(x \wedge w \to y \vee z) \leftrightarrow (y \vee x \vee \neg(w \wedge z)).$$

You do not need to unwind the definition of $\leftrightarrow$, but regard it as an operation with its own truth table.

2.2.17 Exercise. Give an expression a which has the following truth table.

x	y	z	a
0	0	0	0
0	0	1	0
0	1	0	1
0	1	1	1
1	0	0	1
1	0	1	0
1	1	0	0
1	1	1	1

2.2.18 Exercise. Simplify the following expressions using Boolean algebra.

a) $x \wedge (x \to y)$

b) $\neg x \to x$

c) $(x \wedge \neg x) \to y$

d) $x \vee y \to \neg x \wedge y$

2.2.19 Exercise. Show that the inverse is unique; that is, that if $x \wedge y = 0$ and $x \vee y = 1$, then $y = \neg x$.

Hint. Solve the equation $y = \neg x$ using the standard methods.

2.2.20 Exercise (from the exam on 2008-01-09)**.** Simplify $(x \vee y) \to (\neg x \wedge y)$.

2.2.21 Exercise (from the exam on 2007-10-18)**.** Simplify: $(y \vee x) \wedge (x \to y)$.

2.2.22 Exercise (from the exam on 2008-01-09)**.** Solve the equation $x \wedge (y \to z) = (x \wedge y) \to (x \wedge z)$.

2.2.23 Exercise (from the exam on 2007-10-18)**.** Solve the equation $\neg(x \wedge y) \wedge (\neg x \vee z) = \neg(y \wedge z) \to z$.

2.3 Summary

You have learnt how to solve equations in Boolean algebras. We have also introduced implication, which will be important during the whole course. We will not go further into Boolean algebra. The ability to see when two Boolean equations are equal (by solving equations!) is important to decide under which conditions formulas have the same truth value. But it is more important, for the rest of this course, to have the ability to compute algebraically using Boolean algebra.

Chapter 3

Inductively defined sets

3.1 Need for a simple set theory

Anyone who questions the validity of a mathematical result is referred to the proof. He or she must find a weakness in the proof for his or her objection to be taken seriously. The proof is constructed purely on logical steps, so to be able to rely on the proof, one has first to rely on the logic applied. How does one know that it is sound? One method is to use mathematics to prove that it is, but then one runs into an uncomfortable circular reasoning: one justifies mathematics saying that it is correct by the logical laws, which are in turn justified by saying that we have mathematically proven they are correct. Imagine that if the logic is wrong, it will let us do wrong mathematics! In that case one perhaps can, using wrong mathematics, prove that the logic is correct, even if it is not.

Few logicians today believe that one can do logic fully without mathematical methods. Therefore, it is doubtful whether logic in itself could be used as a foundation of mathematics entirely. However, one can settle for a small amount of mathematics when one studies logic, and then the logic can be applied to check more advanced mathematics. The mathematics one needs to do basic logic is a very simple kind of set theory. It is "simple" not in the sense that it should be very easy to understand, but in the sense that it does not have to be powerful enough to contain all mathematics – it is sufficient that it gives us tools for handling what we need to do logic.

The sets we need are all *inductively defined*. That means that one can handle them in such an easy way that it ressembles manipulations in a programming language. In fact, functional programming languages such as OCaml has support for inductively defined sets (which are called *inductive data types*). The most well known inductively defined set is the set of the natural numbers, which is therefore a good first example.

If you learn the principles for inductively defined sets, you will understand the rest more easily, since everything we are going to do follows from these principles.

3.2 Natural numbers

0, 1, 2, 3 and so on are called *natural numbers*. From that explanation we can infer *which* mathematical objects are denoted by "natural numbers" but does it actually say *what* they really are? For example, does it answer the question of what the natural number 3 really is? A moment of thought reminds us that the natural number 3, for instance, could be a number of things (the property of *being three things*) or a point in the number line, or many other things. It is difficult to find an explanation which answers in an exhaustive way *what* the number 3 is.

One can then assert that the different areas of applications have something in common, which leads us to use the label "natural numbers". This common thing reveals itself already in the fact that we can explain which are the natural numbers by saying something as simple as "0, 1, 2, 3, and so on". According to this explanation, the essential property of the number 3 is that it is the successor of the number 2, which is in turn the successor of the number 1, which in turn is the successor of the number 0. We say nothing about what the number 3 can be used for or in which contexts it occurs. The only thing we communicate is the *counting principle*, namely, that one can get the natural numbers through

- starting,
- continuing.

These two points say the essential about counting using natural numbers. We use the natural number 3 for a certain number, since we can count "one, two, three" when we have to deal with three things. In the same way, we use natural numbers for certain points on the number line, since these points can be constructed through the principle of starting and continuing: one chooses a beginning, *origin*, and then continues by pointing into the line in equally long distances, one after another, in a given direction.

Origin: from the latin *origo*, beginning.

This simple explanation of what natural numbers are, that we have just given, shows why we, with our limited mental capabilities, can succeed in handling this infinitely large set of natural numbers: it is not required that we think of infinitely many elements at the same time, but only that we understand the two ways in which natural numbers can be constructed. The whole set of natural numbers is given (induced) by these two principles, so one says that it is *inductively generated* by them. That it is "infinite" only means that there is no limit regarding for how long one can continue generating natural numbers.

We have already the foundations of a theory of inductive definitions. We have created natural numbers and will later, similarly, create other sets which are needed in logic. To do so in an orderly way, we need a better notation than the one we used above. Let us, therefore, reformulate the two rules for creating a natural number:

$$\frac{}{0 \in \mathbb{N}} \qquad \frac{n \in \mathbb{N}}{s(n) \in \mathbb{N}} \,. \tag{3.2.1}$$

This already looks more cryptic, but it is nothing else than a way of writing the two rules "one begins" and "one continues". Every rule is symbolized by a horizontal line. Above each rule we can see what one needs to be allowed to apply it. Under the line we can see what one is allowed to conclude by the help of this rule. The rule on the left has nothing above the line, so one does not need anything to be able to apply it. It says that one is allowed to conclude that "0 is a natural number". In other words: it says that "one begins". The rule on the right says that if one has a natural number, called n, one is allowed to get a new natural number, called $s(n)$. Thus, one only says that for every natural number n, there is a successor. In other words: the rules says "one continues".

The symbol s is traditionally used and stands for "successor".

The two rules can, in the programming language OCaml be applied to define a data type of natural numbers:

```
type n =
  | O
  | S of n
```

These two rules tell us everything about what natural numbers are, if we mean by them the string of words used to count. In contrast, they say very little about what *properties* natural numbers have. For example, it is still difficult to comprehend why Fermat's last theorem is true, even though we know precisely what objects it is about. Nor does the definition say anything about how one can *use* the natural numbers. The connection to numeration, to points on the number line, etc., is not reached at all by the definition. Furthermore, we usually count "0, 1, 2, 3, and so on", rather than "0, $s(0)$, $s(s(0))$, $s(s(s(0)))$, and so on". This is done, of course, because it is confusing to say "the successor of the successor of the successor of zero". The point is that the idea of natural numbers has to do with the fact that we *use* successors, but not necessarily that we *call* them "successors" in our everyday language. In fact, we have to explain for those who are learning to count and read digits that, for instance, 4 is the successor of 3. That is, we must define:

$$\begin{aligned}
1 &\stackrel{\text{def}}{=} s(0)\\
2 &\stackrel{\text{def}}{=} s(1)\\
3 &\stackrel{\text{def}}{=} s(2)\\
4 &\stackrel{\text{def}}{=} s(3)\\
5 &\stackrel{\text{def}}{=} s(4)\\
6 &\stackrel{\text{def}}{=} s(5)\\
7 &\stackrel{\text{def}}{=} s(6)\\
8 &\stackrel{\text{def}}{=} s(7)\\
9 &\stackrel{\text{def}}{=} s(8)
\end{aligned}$$

and then we must also explain that one uses two digits to denote the successor of 9. We leave this for now, since it does not have any relation with inductively defined sets, but rather to how we, in our culture, denote numbers.

We now proceed to define functions on natural numbers. Not even for this any advanced set theory is needed. We have already encountered s, which can be looked as a function of $\mathbb{N}$ into itself. It acts by associating, for every number, its successor. We now define a function which shall *decrease* the value of number one step. This is not possible to do for number 0, as it is "the beginning", so there is no smaller number in the set $\mathbb{N}$. We therefore let the decreasing function be 0 on the number 0.

Now, how does one define a function on $\mathbb{N}$? We shall define it for every element, and there are precisely two sorts of elements, since there are two rules to form elements in $\mathbb{N}$: a natural numebr is either of the form 0 or of the form $s(n)$, where n is a natural number. We call the function p and define it on both sorts:

p is used for Predecessor.

$$p(0) \stackrel{\text{def}}{=} 0 \tag{3.2.2}$$

$$p(s(n)) \stackrel{\text{def}}{=} n. \tag{3.2.3}$$

We therefore define the function by saying how it is computed for all sorts of elements in the set.

In OCaml one can write the definition as follows:

```
let p = function
  | O -> O
  | S x -> x
```

Addition is defined in a similar manner, in two cases:

$$a + 0 \stackrel{\text{def}}{=} a \tag{3.2.4}$$

$$a + s(n) \stackrel{\text{def}}{=} s(a + n) \tag{3.2.5}$$

Addition occurs on the right side of (3.2.5). One therefore says that the definition is "recursive". The computation of $3 + 2$ is done as follows, by unwinding definitions:

Here we use the definition of +, 2 and 1.

$$3 + 2 \stackrel{\text{def}}{=} 3 + s(1) \stackrel{\text{def}}{=} s(3 + 1) \stackrel{\text{def}}{=} s(3 + s(0)) \stackrel{\text{def}}{=} s(s(3 + 0)) \stackrel{\text{def}}{=} s(s(3)). \tag{3.2.6}$$

We can now use the definition of 4 and 5 reversely, so that we get $s(s(3)) \stackrel{\text{def}}{=} s(4) \stackrel{\text{def}}{=} 5$ and we can give the answer "5" as the result of this computation.

In OCaml one must specify that + occurs on the right hand side by writing `rec` for "recursive" in the definition:

```
let rec plus a = function
  | O -> a
  | S x -> S (plus a x)
```

The definition of brings nothing new, except that one can choose to have the recursion in the first argument, if one wants to:

$$0 \cdot a \stackrel{\text{def}}{=} 0 \tag{3.2.7}$$
$$s(n) \cdot a \stackrel{\text{def}}{=} (n \cdot a) + a \tag{3.2.8}$$

In the definition of exponentiation, the recursion has to occur in the second argument:

$$a^0 \stackrel{\text{def}}{=} 1 \tag{3.2.9}$$
$$a^{s(n)} \stackrel{\text{def}}{=} a^n \cdot a. \tag{3.2.10}$$

We can now compute in a completely automatic way, by unwinding definitions,

$$2^3 \stackrel{\text{def}}{=} 2^{s(2)} \tag{3.2.11}$$
$$\stackrel{\text{def}}{=} 2^2 \cdot 2 \tag{3.2.12}$$
$$\stackrel{\text{def}}{=} 2^{s(1)} \cdot 2 \tag{3.2.13}$$
$$\stackrel{\text{def}}{=} (2^1 \cdot 2) \cdot 2 \tag{3.2.14}$$
$$\stackrel{\text{def}}{=} (2^{s(0)} \cdot 2) \cdot 2 \tag{3.2.15}$$
$$\stackrel{\text{def}}{=} ((2^0 \cdot 2) \cdot 2) \cdot 2 \tag{3.2.16}$$
$$\stackrel{\text{def}}{=} ((1 \cdot 2) \cdot 2) \cdot 2 \tag{3.2.17}$$
$$\stackrel{\text{def}}{=} ((s(0) \cdot 2) \cdot 2) \cdot 2 \tag{3.2.18}$$
$$\stackrel{\text{def}}{=} ((0 \cdot 2 + 2) \cdot 2) \cdot 2 \tag{3.2.19}$$
$$\stackrel{\text{def}}{=} ((0 + 2) \cdot 2) \cdot 2 \tag{3.2.20}$$
$$\stackrel{\text{def}}{=} ((0 + s(1)) \cdot 2) \cdot 2 \tag{3.2.21}$$
$$\stackrel{\text{def}}{=} (s(0 + 1) \cdot 2) \cdot 2 \tag{3.2.22}$$
$$\stackrel{\text{def}}{=} (s(0 + s(0)) \cdot 2) \cdot 2 \tag{3.2.23}$$
$$\stackrel{\text{def}}{=} (s(s(0 + 0)) \cdot 2) \cdot 2 \tag{3.2.24}$$
$$\stackrel{\text{def}}{=} (s(s(0)) \cdot 2) \cdot 2 \tag{3.2.25}$$

and so on. It is apparent that this is an extremely time consuming process, but it follows simple principles and the definitions clearly convey the idea of how computation works. For more efficient calculations one has to find smarter ways to compute the result.

For the sake of completeness, we also consider the definition of minus. Since we have no negative numbers, we let $a - b$ be zero in case b is greater than a.

$$a - 0 \stackrel{\text{def}}{=} a \tag{3.2.26}$$
$$a - s(n) \stackrel{\text{def}}{=} p(a - n) \tag{3.2.27}$$

3.2.28 Exercise. Compute $1 + 2$, $1 \cdot 2$ and $1 - 2$ by unwinding the definitions (no short cuts!).

3.2.29 Exercise. Describe what the following function does (defined on natural numbers):

$$f(a, 0) \stackrel{\text{def}}{=} a$$
$$f(a, s(n)) \stackrel{\text{def}}{=} s(f(p(a), n))$$

3.3 The algebra of two elements

We can define the algebra with two elements inductively. We call it Boole.

$$\frac{}{0 \in \text{Boole}} \qquad \frac{}{1 \in \text{Boole}} \tag{3.3.1}$$

The operations are defined using the same principles as for natural numbers. Since the set is defined by two rules, there are two rows when we define functions.

3.3.2 Definition.

$$\begin{aligned} \neg 0 &\stackrel{\text{def}}{=} 1 & a \wedge 0 &\stackrel{\text{def}}{=} 0 & a \vee 0 &\stackrel{\text{def}}{=} a \\ \neg 1 &\stackrel{\text{def}}{=} 0 & a \wedge 1 &\stackrel{\text{def}}{=} a & a \vee 1 &\stackrel{\text{def}}{=} 1 \end{aligned}$$

3.3.3 Exercise. Use Definition 2.2.6 to compute $a \to 0$ and $a \to 1$ if the other operations are defined according to 3.3.2.

3.4 Induction and recursion

We have already seen the principle of *recursion*: when a set is inductively defined, one defines functions from it by giving the function values for the different cases that can occur. That is how we defined predecessor, as well as addition, multiplication, and more, as well as the Boolean operations. An important sophistication of recursion is that the computation of a value can lead to a new expression that in turn has to be computed, and which in itself contains the function which is to be calculated. For instance, if one tries to compute $a+s(b)$, one gets the expression $s(a+b)$, which itself contains $+$. One has to be careful not to give definitions which lead to infinite computations. For example, it is not allowed to define a function f as:

$$f(0) \stackrel{\text{def}}{=} 0 \tag{3.4.1}$$

$$f(s(n)) \stackrel{\text{def}}{=} f(s(s(n))) \tag{3.4.2}$$

since computing $f(1)$ would lead to a series of computation steps which never end. It is difficult to be precise as to which recursive definitions are acceptable and which must be avoided. This is a question better reserved for deeper studies on inductively defined sets. We shall not cover that in the course, but will content ourselves to verify manually that each computation in a recursive definition terminates after a finite number of steps.

Related to recursion is *induction*. While recursion is used for defining functions, induction is used to prove things about inductively defined sets. These are two different tools, but the way of working with them are so alike that is very easy to confuse them.

The idea in both recursion and induction is the following. Inductively defined sets are described by a number of rules. Every element in an inductively defined set is therefore in one of the forms which occurs underneath these rules lines. For example, natural numbers are either of the form 0 or $s(n)$. If one wants to define a function on all natural numbers, or prove a theorem about all natural numebrs, it is sufficient to consider numbers of each of these forms. Furthermore, one can assume that the function values, or the validity of the theorems, are already established for the parts involved. If, for example, one defines a function on the natural numbers and considers the case $s(n)$, one assumes that the function value on n is already given. If one, in a similar fashion, proves a theorem about natural numbers and considers the case $s(n)$, one assumes that the validity of the theorem is already given in the case n. These hypotheses, the *inductive hypothesis*, are justified by the fact that the set is built up only by these given rules. Therefore, one only needs to check that the validity of the theorem is preserved in every step of this process.

We will consider some examples of proof by induction.

3.4.3 Theorem. *For every natural number x we have $0 + x = x$.*

Proof. We shall prove the theorem in two cases, since there are two ways of constructing natural numbers (3.2.1).

1. x is of the form 0. In this case, we need to show that $0 + 0 = 0$, but this follows directly from the definition of + (3.2.4). This case is, thus, clear.

2. x is of the form $s(n)$. In this case we need to show that $0 + s(n) = s(n)$. we use (3.2.5) and see that the left hand side is transformed into $s(0+n)$. The inductive hypothesis says that $0 + n = n$, so $s(0 + n) = s(n)$. Then this case is also covered.

□

Not only with natural numbers can one do proofs by induction. It works for all inductively defined sets. In some cases, such as the case Boole, no inductive hypothesis occur, since the rules (3.3.1) have nothing above the line. The following theorem illustrate this:

3.4.4 Theorem. *In the set* Boole*, for every element x we have $x \wedge x = x$.*

Proof. Since Boole is defined by two rules (3.3.1) there are two cases to check.

1. x is of the form 0. In this case we need to show that $0 \wedge 0 = 0$. But it follows from the definition that $a \wedge 0 \stackrel{\text{def}}{=} 0$.

2. x is fo the form 1. In this case we need to show that $1 \wedge 1 = 1$. But it follows from the definition that $a \wedge 1 \stackrel{\text{def}}{=} a$.

□

Later, when we define more complicated sets inductively, we shall have more complicated induction proofs – more cases and more inductive hypothesis. The principles are always the same: for every rule we have, defining the set, one gets a case to consider, and one gets an inductive hypothesis for every expression above this rule's line.

3.5 Summary

You have seen how one can introduce and reason about *inductively defined sets* without referring to more advanced set theory or more advanced results in mathematics. An example is the set of natural numbers. You have seen how the foundations of arithmetic is built according to these principles and using recursion. It has been explained why inductively generated sets lead naturally to induction proofs and recursion, as well as how these tools are applied. The most important for you to take into account for the rest of this course is precisely the insight of how to define sets inductively and perform inductive proofs, as well as how to define functions using recursion.

Part II

Propositional logic

Chapter 4

The language and semantics of propositional logic

4.1 Logical formulas

This course is about mathematical logic, which means that one handles propositions in a mathematical way. There is no more mystery in this than there is in mathematics about natural numbers; we just have a set of propositions rather than a set of numbers to work with. We simply change the definition of natural number a little and we will get propositions rather than numbers.

To begin with, we need *propositional variables*; that is, symbols for arbitrary propositions. We shall dedicate ourselves to *formal* logic, which is called like that precisely because the logical rules should not be affected by what the propositions mean – only the *form* shall be relevant. For instance, we will be able to say things such as: if the propositions P_1 and P_2 are true, then the proposition $P_1 \wedge P_2$ is true – and this will hold no matter what P_1 and P_2 stand for. That is why one denotes them by non descriptive letters: to stress the fact that their meaning can vary. It will not be sufficient with P_1 and P_2, there is no limit to how many propositional variables we may need. Instead of deciding the number of variables in advance, we leave it open and say it in the following way: the set Pvar of propositional variables is defined inductively through the following rules, where n counts how many propositional variables we want:

$$\frac{}{P_1 \in \mathrm{Pvar}} \quad \cdots \quad \frac{}{P_n \in \mathrm{Pvar}} \tag{4.1.1}$$

We therefore have n such rules, and each and everyone of them says that something is a propositional variable. This is not entirely satisfactory, since we would also want to have composed propositions such as $P_1 \wedge P_2$. We shall call them *formulas* rather than propositions, since they are just formal expressions, whose meaning depend on how one interprets P_1 and P_2. More specifically, we define the set Form inductively by the following rules, where Greek letters are used as variables in Form in the same manner as we use Latin letters for variables in $\mathbb{N}$.

Metavariables: $\varphi, \psi, \ldots$
Object variables: $P_1, P_2, \ldots$

On one hand, Greek variables are used as variables; on the other hand, $P_1, \ldots, P_n$ will denote propositional variables. These are two different sort of variables.

The difference is that $\varphi, \psi, \ldots$ stand for *arbitrary formulas*, while $P_1, P_2, \ldots$ are *specific formulas* which one can *think* of as variables, namely, by imagining that they stand for propositions such as “the sun is shining” or “the grass is green”.

Compare this with the set of polynomials- with real coefficients-. In this set, x is a *specific* element, namely the polynomial x. we say such things as “let p be an arbitrary polynomial”, but we cannot say “let x be an arbitrary polynomial”. Here p varies over the set of polynomials (p is a metavariable), while x is fixed as symbolizing a variable (x is an object variable).

In practice, this difference means that we can say things such as “let φ be any formula” while the same meaning should be nonsensical if we exchange φ for P_1, which could never be any formula – since P_1 is a specific fixed formula.

> Note that the set Form depends on n. We could write Form(n) instead of Form, but in practice we will not have any difficulties if we omit n.

4.1.2 Definition.

$$\frac{}{P_1 \in \text{Form}} \quad \cdots \quad \frac{}{P_n \in \text{Form}}$$

$$\frac{}{\top \in \text{Form}}$$

$$\frac{}{\bot \in \text{Form}}$$

$$\frac{\varphi \in \text{Form} \qquad \psi \in \text{Form}}{(\varphi \wedge \psi) \in \text{Form}}$$

$$\frac{\varphi \in \text{Form} \qquad \psi \in \text{Form}}{(\varphi \vee \psi) \in \text{Form}}$$

$$\frac{\varphi \in \text{Form} \qquad \psi \in \text{Form}}{(\varphi \to \psi) \in \text{Form}}$$

The formula $\top$ should be thought of as a proposition which is always true, while $\bot$ symbolizes a proposition which is always false. You may miss $\neg$ from Boolean algebras. We will omit this operation, since it is cumbersome to handle too many of them. Instead, we will look at $\neg\varphi$ as an abbreviation of $\varphi \to \bot$ (compare Exercises 2.2.12 and 3.3.3).

4.1.3 Definition. We introduce the following abbreviations, if φ, ψ are formulas:

$$\neg\varphi \overset{\text{def}}{=} (\varphi \to \bot)$$
$$\varphi \leftrightarrow \psi \overset{\text{def}}{=} ((\varphi \to \psi) \wedge (\psi \to \varphi))$$

> Since $\wedge$, $\vee$, $\to$ are used to join formulas together into bigger formulas, they are called *connectives*. We will even call $\top$ and $\bot$ connectives, more specifically *nullary connectives*.

If we want to write the proposition "the sun is shining and the grass is green" as an element in Form we can let the variable P_1 stand for "the sun is shining" and the variable P_2 for "the grass is green", so that we can express what we want with the formula $P_1 \wedge P_2$, which is an element of Form. We must be careful if we have to express "the sun is shining, the grass is green and I am happy" as an element in Form, since there are actually two formulas which can express this. Spontaneously one might want to say $P_1 \wedge P_2 \wedge P_3$, but we have not introduced the possibility of constructing ternary $\wedge$-propositions, only binaries. We must therefore differentiate the formulas constructed in the following way:

$$\frac{\frac{}{P_1 \in \text{Form}} \qquad \frac{\frac{}{P_2 \in \text{Form}} \quad \frac{}{P_3 \in \text{Form}}}{(P_2 \wedge P_3) \in \text{Form}}}{(P_1 \wedge (P_2 \wedge P_3)) \in \text{Form}} \tag{4.1.4}$$

respectively

$$\frac{\frac{\frac{}{P_1 \in \text{Form}} \quad \frac{}{P_2 \in \text{Form}}}{(P_1 \wedge P_2) \in \text{Form}} \qquad \frac{}{P_3 \in \text{Form}}}{((P_1 \wedge P_2) \wedge P_3) \in \text{Form}\,.} \tag{4.1.5}$$

Note that parentheses can be used in the end formula to make it clear how it is formed. Sometimes it is superfluous to know how a proposition was formed, and then one can disregard parentheses, but sometimes they are necessary, for example, $P_2 \wedge P_3$ is a "subformula" of $P_1 \wedge (P_2 \wedge P_3)$, but it is not a subformula of $(P_1 \wedge P_2) \wedge P_3$, so when we later deal with subformulas, it will be impossible to be careless and write "$P_1 \wedge P_2 \wedge P_3$". In other contexts, it will often not be a problem

> Parentheses are used to say in which order one constructs a formula.
>
> - sometimes it is important to put parentheses,
> - sometimes they are not needed.

In the definition, we have not said a word about the fact that we will let $\wedge$ symbolize "and", and so on. Precisely as in the case of the natural numbers, whose meaning could be numeration, points on a number line or something else, we will not mention in the definition itself what we use the formulas for. Until further notice, they will just be empty formal expressions, whose meaning we can decide upon later.

It is important to remember the difference between what we have just done and Boolean algebra. In Boolean algebras we have, for instance, $a \wedge a = a$, but this is not true in the set Form of formulas. The formula $P_1 \wedge P_1$ is a different formula than P_1. The equality $=$ here means *the same formula*, not *the same value*. The set Form is simply not a Boolean algebra, the rules of computation do not hold.

The set Form is not a Boolean algebra, even though the notation is very similar!

4.1.6 Exercise. Derive, as in (4.1.4) and (4.1.5):

a) $((P_1 \to P_2) \wedge \bot) \in \text{Form}$

b) $\neg P_1 \in \text{Form}$

c) $\bot \leftrightarrow \top \in \text{Form}$

4.2 Semantics

In the previous section we introduced the set Form of propositional formulas, but mentioned only vaguely how the formulas should be interpreted. In this section we shall formulate it more mathematically.

First of all, we say that the basic formulas $P_1, \dots, P_n$ can be interpreted in many ways. One can think of them as "the sun is shining" or "the grass is green", but also as mathematical propositions like "3 is a prime number" – and even as false mathematical propositions: "all prime numbers are odd". When we think of the formulas $P_1, \dots, P_n$ as specific propositions, we say that we give an *interpretation* of the formulas. The formulas are the same, but the interpretations can vary in infinitely many ways. Often one denotes an interpretation by $\mathcal{A}$. One says that the formulas $P_1, \dots, P_n$ are interpreted as the propositions $P_1^{\mathcal{A}}, \dots, P_n^{\mathcal{A}}$. An interpretation is thus a sort of function, which for every propositional variable P_i assigns a proposition $P_i^{\mathcal{A}}$. Since the propositional variables are interpreted, we automatically get an interpretation of all formulas by reading $\wedge$ as *and*, $\vee$ as *or*, $\to$ as *entails* (or as *if ... then* ...), $\top$ as *the true* and $\bot$ as *the false*. As pointed out when we were doing Boolean algebra, these expressions from natural language are often ambiguous. Therefore, one uses *truth values* to be more precise. The truth value of a formula is 1 if it is interpreted as a true proposition, and 0 if it is interpreted as a false proposition. One often denotes truth values by double square brackets. Since this is a function on Form, which is inductively defined, the definition becomes recursive.

P_1 is a formula, $P_1^{\mathcal{A}}$ is a proposition

4.2.1 Definition. The truth value of a formula is an element in Boole determined by an interpretation $\mathcal{A}$ through:

$$[\![P_i]\!] \stackrel{\text{def}}{=} \begin{cases} 1 & \text{if } P_i^{\mathcal{A}} \text{ is true} \\ 0 & \text{if } P_i^{\mathcal{A}} \text{ is false} \end{cases} \tag{4.2.2}$$

$$[\![\top]\!] \stackrel{\text{def}}{=} 1 \tag{4.2.3}$$

$$[\![\bot]\!] \stackrel{\text{def}}{=} 0 \tag{4.2.4}$$

$$[\![\varphi \wedge \psi]\!] \stackrel{\text{def}}{=} [\![\varphi]\!] \wedge [\![\psi]\!] \tag{4.2.5}$$

$$[\![\varphi \vee \psi]\!] \stackrel{\text{def}}{=} [\![\varphi]\!] \vee [\![\psi]\!] \tag{4.2.6}$$

$$[\![\varphi \to \psi]\!] \stackrel{\text{def}}{=} [\![\varphi]\!] \to [\![\psi]\!] \tag{4.2.7}$$

Since $[\![\varphi]\!]$ depends on the interpretation $\mathcal{A}$, one often writes $[\![\varphi]\!]^{\mathcal{A}}$. When there is no possibility of confusion, though, one only writes $[\![\varphi]\!]$.

4.2.8 Exercise. Check that the following holds for all formulas φ, ψ.

a) $[\![\neg\varphi]\!] = \neg[\![\varphi]\!]$.

b) $[\![\varphi \leftrightarrow \psi]\!] = [\![\varphi]\!] \leftrightarrow [\![\psi]\!]$.

Hint. It may look obvious, but the problem is that we have defined the operation differently in Boolean algebras and in Form. Use the definitions 4.1.3 and 2.2.14, as well as Exercise 3.3.3.

4.2.9 Example. Assume that $P_1^{\mathcal{A}}$ and $P_2^{\mathcal{A}}$ are true. Compute $[\![\neg(P_1 \vee P_2) \to \bot]\!]$.

Solution.

$$\begin{aligned}
[\![\neg(P_1 \vee P_2) \to \bot]\!] &\stackrel{\text{def}}{=} [\![\neg(P_1 \vee P_2)]\!] \to [\![\bot]\!] && (4.2.10)\\
&= \neg[\![P_1 \vee P_2]\!] \to [\![\bot]\!] && (4.2.11)\\
&\stackrel{\text{def}}{=} \neg([\![P_1]\!] \vee [\![P_2]\!]) \to [\![\bot]\!] && (4.2.12)\\
&\stackrel{\text{def}}{=} \neg(1 \vee 1) \to 0 && (4.2.13)\\
&\stackrel{\text{def}}{=} \neg 1 \to 0 && (4.2.14)\\
&\stackrel{\text{def}}{=} 0 \to 0 && (4.2.15)\\
&\stackrel{\text{def}}{=} 1 && (4.2.16)
\end{aligned}$$

□

The word *tautology* is explained in dictionaries as "unnecessary repetition", but it is not in this sense that it is used in logic, although the etymology is the same. In antiquity, propositions of the type "humans are humans" or "odd numbers are odd" used to be considered. Here, there is certainly a needless repetition, but logicians have instead focused on the obviousness of these propositions. Already Fredrik Afzelius[a] (1839) wrote: "Analytical judgements, in which the predicates as well as its characteristics, are not obviously contained in the subject, but which completely coincide with themselves are called tautological; for instance, humans are humans."[b] Etymologically, the prefix "tauto" has the same origin as "auto" and means *self*, while "logy" comes from the greek "logos" and can mean both *word* and *reason*, *rationality*, *clarity*. One can thus translate "tautology" directly as "self word" (repetition) or "self-evident".

[a] Afzelius (1812-1896) was a teacher of philosophy at Uppsala University, and an early exponent of Hegel's ideas in Sweden.

[b] Suggested translation of "Analytiska Omdömen, i hvilka Predikatet icke blott inhålles i Subjektet såsom dess Kä nnetecken, utan alldeles sammanfaller kallas Tavtologiska, t. ex. Menniskan är Menniska."

4.2.17 Exercise. Assume that $P_1^{\mathcal{A}}$, $P_2^{\mathcal{A}}$, $P_4^{\mathcal{A}}$ are false, while $P_3^{\mathcal{A}}$ and $P_5^{\mathcal{A}}$ are true. Compute $[\![\neg(P_2 \to \neg P_3) \wedge (P_1 \to P_5)]\!]$. Justify!

4.2.18 Example. Compute $[\![(P_1 \vee \neg(P_2 \wedge P_3)) \to \top]\!]$.

Solution. Here we do not need to know whether $P_1^{\mathcal{A}}$, $P_2^{\mathcal{A}}$, $P_3^{\mathcal{A}}$ are true, since Exercise 3.3.3 allows us to go almost directly to the answer.

$$\begin{aligned}
[\![(P_1 \vee \neg(P_2 \wedge P_3)) \to \top]\!] &= [\![P_1 \vee \neg(P_2 \wedge P_3)]\!] \to [\![\top]\!] && (4.2.19)\\
&= [\![P_1 \vee \neg(P_2 \wedge P_3)]\!] \to 1 && (4.2.20)\\
&= 1 && (4.2.21)
\end{aligned}$$

□

With some experience, one sees that while calculating truth values, the recursion makes the square brackets "move in" into the smaller parts of the formula. One can therefore skip several steps.

Solution Example 4.2.9. $[\![\neg(P_1 \vee P_2) \to \bot]\!] = \neg(1 \vee 1) \to 0 \stackrel{\text{def}}{=} 0 \to 0 \stackrel{\text{def}}{=} 1$. □

4.2.22 Example. Compute $[\![(P_1 \to P_2) \vee (P_2 \to P_1)]\!]$.

Solution.

$$\begin{aligned}
[\![(P_1 \to P_2) \vee (P_2 \to P_1)]\!] &\stackrel{\text{def}}{=} ([\![P_1]\!] \to [\![P_2]\!]) \vee ([\![P_2]\!] \to [\![P_1]\!]) && (4.2.23)\\
&= \neg[\![P_1]\!] \vee [\![P_2]\!] \vee \neg[\![P_2]\!] \vee [\![P_1]\!] && (4.2.24)\\
&= 1 && (4.2.25)
\end{aligned}$$

□

Here we have an example where one does not need to know if the variables involved are interpreted as true formulas or not. The formula has value 1 in all interpretations – we say that it is "true in all interpretations". Such a formula is called a *tautology*.

▶ **4.2.26 Definition.** A *tautology* is a formula which is true in all interpretations.

If one would like to investigate if a formula is a tautology, one could check, using Boolean algebra, if its truth value is 1 no matter what the values of the propositional variables are, as we did in the previous example. Another way is to construct a truth table for the formula.

4.2.27 Example. Check using a truth table that $(P_1 \to P_2) \lor (P_2 \to P_1)$ is a tautology.

Solution. Below P_1 and P_2 we write the corresponding truth values in all possible combinations. In general we write the table as usual:

(	P_1	$\to$	P_2	)	$\lor$	(	P_2	$\to$	P_1	)
	0	1	0		**1**		0	1	0	
	0	1	1		**1**		1	0	0	
	1	0	0		**1**		0	1	1	
	1	1	1		**1**		1	1	1	

(4.2.28)

Answer: since there are only ones in the rows underneath $\lor$, the formula is a tautology. □

If one chooses to compute with Boolean algebras, it is convenient to use the following definition

When one calculates using $\approx$ one does not need to put so many parentheses. The fact that $P_1 \lor P_2 \lor P_3$ can be interpreted as two different formulas does not matter, since their value is the same.

▶ **4.2.29 Definition** (logical equivalence)**.** If φ, ψ are formulas, then $\varphi \approx \psi$ means that $[\![\varphi]\!] = [\![\psi]\!]$ is true in *all* interpretations.

4.2.30 Example. Check, using Boolean algebras, that $(P_1 \to P_2) \lor (P_2 \to P_1)$ is a tautology.

Solution. $(P_1 \to P_2) \lor (P_2 \to P_1) \approx \neg P_1 \lor P_2 \lor \neg P_2 \lor P_1 \approx \top$. □

4.2.31 Exercise. Show that if φ, ψ are formulas, then $\varphi \leftrightarrow \psi$ is a tautology if and only if $\varphi \approx \psi$.

A difference between $=$ and $\approx$:

$$P_1 \land P_1 \neq P_1$$
$$P_1 \land P_1 \approx P_1$$

4.2.32 Exercise. Decide which of the following formulas are tautologies.

a) $\neg(P_1 \land P_2) \leftrightarrow (P_1 \to \neg P_2)$

b) $(P_1 \to P_2) \lor (P_2 \to P_3)$

c) $(P_1 \to (P_2 \to P_3)) \leftrightarrow ((P_1 \land P_2) \to P_3)$

d) $((P_1 \land P_4) \to (P_2 \lor P_3)) \leftrightarrow (\neg P_1 \lor P_2 \lor P_3 \lor P_4)$

Hint. It is often convenient to use Exercise 4.2.31.

4.2.33 Exercise (from the exam on 2003-01-09)**.** Decide whether the following formula is a tautology:

$$(P_3 \lor P_1 \to \neg P_2 \land P_3) \to (P_2 \lor \neg P_3 \to \neg P_1)\,.$$

4.2.34 Exercise (from the exam on 2002-08-20)**.** Assume that φ, ψ, σ are formulas that satisfy:

$$\varphi \land \sigma \approx \psi \land \sigma$$
$$\varphi \lor \sigma \approx \psi \lor \sigma\,.$$

Show that $\varphi \approx \psi$.

4.2.35 Exercise (from the exam on 2004-10-18)**.** Decide whether $((P_2 \to P_1) \to P_2) \to P_2$ is a tautology.

4.2.36 Exercise (from the exam on 2005-01-07)**.** Decide whether $((P_1 \to P_1) \to P_1) \to P_1$ is a tautology.

4.2.37 Exercise (from the exam on 2005-08-23)**.** Is

$$((P_1 \land P_2) \to P_3) \leftrightarrow ((P_1 \to P_3) \lor (P_2 \to P_3))$$

a tautology?

4.2.38 Exercise (from the exam on 2004-08-17)**.** Decide whether $(P_3 \to \neg P_1) \lor \neg P_2$ is a tautology.

The term *model* was coined when Felix Klein (1849–1925) constructed a "model" of geometry with the negation of the parallel axiom; that is, a new interpretation of the geometric concepts in which the parallel axiom is false.

Do you understand the difference between $\mathcal{A} \models \varphi$ and $\psi \models \varphi$?

▶ **4.2.39 Definition.** By $\mathcal{A} \models \varphi$ it is meant that $[\![\varphi]\!]^{\mathcal{A}} = 1$. In that case one says that "φ is true in $\mathcal{A}$" or that "$\mathcal{A}$ is a model of φ". If Γ is a set of formulas, "$\mathcal{A}$ is a model of Γ" means that $\mathcal{A}$ is a model of *every* formula in Γ.

▶ **4.2.40 Definition.** By $\varphi_1, \ldots, \varphi_n \models \varphi$ it is meant that φ is true in every interpretation in which $\varphi_1, \ldots, \varphi_n$ are true. One says that φ is a *logical consequence* of $\varphi_1, \ldots, \varphi_n$.

4.2.41 Exercise. What does one get from Definition 4.2.40 when $n = 1$ respectively $n = 0$?

4.3 Summary

You have seen how the set Form is defined and how one defines important functions on it by recursion. An example of such a function is the truth value function $[\![\cdot]\!]$. You have encountered the term *interpretation* and have practiced deciding whether a formula is a *tautology*, that is, whether it is true in every interpretation. You have seen how one can use truth tables for this, as well as algebraic methods for Boolean algebras. The most important thing to take with you for the rest of the course is the understanding of the notions of *interpretation* and *tautology* as well as an insight of what the set Form consists of. The notation $[\![\varphi]\!]^{\mathcal{A}}$ will be used a lot, so you should make sure to have a good understanding of it.

Chapter 5

Natural deduction

In the previous chapter we defined how truth values for formulas are calculated in various interpretations. We shall now forget the interpretations for a while and think about logic in yet another way. In the next chapter we shall prove that what we do here is in fact "sound" regarding the semantics (Soundness theorem 6.1.5, page 45).

Now we shall instead approach the formulas reminding ourselves how one usually reasons about *and*, *or* and *if ... then ...* and try to expose such rules with a horizontal line as we have previously seen. We shall discuss some different rules, and then, little by little, a limited set of rules will emerge. The final rules are collected in Figure 5.1 on page 40. It is important that you learn these rules carefully. They have to be used precisely as they stand, with φ, ψ, σ substituted by arbitrary formulas. What the various dots and square brackets mean will be explained soon.

We construct derivations which graphically look like a tree. It is often easier to construct them from below, but they are best read from top to bottom. When one speaks or writes about a derivation, one expresses oneself as if what is below comes *after* what is above, even though the derivation is constructed in the other way around.

5.1 Conjunction

What does one have to know to conclude that the formula $\varphi \wedge \psi$ symbolizes a true proposition? The typical situation is that one knows that φ and ψ symbolize true propositions. We therefore establish the following rule:

$$\frac{\varphi \text{ true} \qquad \psi \text{ true}}{\varphi \wedge \psi \text{ true}}\,. \tag{5.1.1}$$

For the sake of simplicity, we will skip, in what follows, the label "true". We will write instead

$$\frac{\varphi \qquad \psi}{\varphi \wedge \psi}\,. \tag{5.1.2}$$

This is not the only rule we need for $\wedge$. With it we can only *introduce* conjunctions, but we can never get rid of them. We therefore call this rule an *introduction rule* and we also introduce two *elimination rules*

$$\frac{\varphi \wedge \psi}{\varphi} \qquad\qquad \frac{\varphi \wedge \psi}{\psi} \tag{5.1.3}$$

which says that if $\varphi \wedge \psi$ symbolize a true proposition, then so does φ, as well as ψ.

We might as well establish rules for quinary conjunction:

$$\frac{\varphi_1 \qquad \varphi_2 \qquad \varphi_3 \qquad \varphi_4 \qquad \varphi_5}{\varphi_1 \wedge \varphi_2 \wedge \varphi_3 \wedge \varphi_4 \wedge \varphi_5}\,, \tag{5.1.4}$$

$$\frac{\varphi_1 \wedge \varphi_2 \wedge \varphi_3 \wedge \varphi_4 \wedge \varphi_5}{\varphi_1} \quad \ldots \quad \frac{\varphi_1 \wedge \varphi_2 \wedge \varphi_3 \wedge \varphi_4 \wedge \varphi_5}{\varphi_5} \tag{5.1.5}$$

but we will just consider binary conjunctions, since those are the ones we have in Form. Although it is worth mentioning that for an n-ary conjunction we will get:

1. one *introduction rule* which consists of n formulas above the line and the n-ary conjunction itself underneath the line,

The conjuncts are formulas which are glued together by applying $\wedge$ to get more complex formulas.

2. n *elimination rules* which consist of the conjunction itself above the line and– in the i-th elimination rule– the i-th conjunct underneath the line.

This perhaps comes useful to us, since we have a 0-ary conjunction in Form – given that we have decided that $\top$ is such a thing. Thus, it should have an introduction rule which has 0 formulas above the line and the formula $\top$ below it:

$$\frac{}{\top} \tag{5.1.6}$$

According to the same analogy, it should have 0 elimination rules. Thus, the only rule we have about $\top$ says that without knowing anything in particular, we can conclude that $\top$ symbolizes the true proposition.

With the rules we have just written down, we can derive in the shape of a tree, for instance, that the formula $(\top\wedge\top)\wedge(\top\wedge\top)$ symbolizes a true formula.

$$\frac{\dfrac{\dfrac{}{\top} \quad \dfrac{}{\top}}{\top\wedge\top} \qquad \dfrac{\dfrac{}{\top} \quad \dfrac{}{\top}}{\top\wedge\top}}{(\top\wedge\top)\wedge(\top\wedge\top)} \tag{5.1.7}$$

We say “derive” when we construct these trees, but we cannot yet be certain that the trees work to prove things. We shall show later that this is the case, but until then you should look at the trees as pure formal mathematical objects that we manipulate.

One calls this way of deriving formulas *natural deduction*, since it ressembles how one reasons informally in mathematics.

One can also, starting with the formula $P_1\wedge P_2$, derive the formula $P_2\wedge P_1$.

$$\frac{\dfrac{P_1\wedge P_2}{P_2} \qquad \dfrac{P_1\wedge P_2}{P_1}}{P_2\wedge P_1} \tag{5.1.8}$$

Observe the difference between (5.1.7) and (5.1.8). In the former derivation, all formulas at the top of the tree have a line above them, which meant that we could conclude that they symbolize true propositions. In the latter tree, there are no such lines, and that is sensible if we consider that $P_1\wedge P_2$ does not always symbolize a true proposition – its truth value depends on how we interpret P_1 and P_2. In the first case (5.1.7) one says that one has “derived $(\top\wedge\top)\wedge(\top\wedge\top)$” and writes

$$\vdash (\top\wedge\top)\wedge(\top\wedge\top) \tag{5.1.9}$$

Remember the difference between $\vDash$ and $\vdash$.

to denote that there exists a derivation of the formula. In the second case (5.1.8) one says that one has “derived $P_2\wedge P_1$ from $P_1\wedge P_2$” and writes

$$(P_1\wedge P_2)\vdash(P_2\wedge P_1) \tag{5.1.10}$$

to denote that there is a derivation from $P_1\wedge P_2$ to $P_2\wedge P_1$. One thinks of $P_1\wedge P_2$ as an *assumption* (sometimes called *hypothesis*), so that (5.1.8) symbolizes the argument.

Assume that $P_1\wedge P_2$ is true. (...) Then $P_2\wedge P_1$ is also true. (5.1.11)

5.2 Implication

When one has done an argument as (5.1.11) in a mathematical proof, one usually summarizes the situation by claiming an implication:

If $P_1\wedge P_2$, then $P_2\wedge P_1$.

Such propositions are what we symbolize with $\to$. From this we find the natural introduction rule for $\to$: if one could derive a formula ψ from a formula φ, then $\varphi \to \psi$ is true. Furthermore: the formula φ was an *assumption* when we derived ψ, but it is no longer used as an assumption when we have concluded that $\varphi \to \psi$ is true. Consider, for example, the following argument:

> Assume that n is odd. (...) Thus, n^2 is odd. Therefore, it is true that if n is odd, then n^2 is odd.

The last sentence says that "if n is odd, then n^2 is odd", it does not say "assume that n is odd, then it is true that if n is odd then n^2 is odd". The assumption that n is odd is only used during the argument, but it is later *discharged*. One marks discharged assumptions by putting them within square brackets. Therefore, the rules of implication introduction become:

$$\dfrac{\begin{matrix}[\varphi] \\ \vdots \\ \psi\end{matrix}}{\varphi \to \psi} \tag{5.2.1}$$

Here we have a big difference compared to *all* the previous rules we have expressed in this line form. In these rules, there were the *formulas* which were above the line the ones one needed to be able to apply the rule. Here, it is the whole *derivation* above the line one has to possess.

While the previous sort of rules are called *inference rules* one often calls this latter form of rules *deduction rules*. This is not something you need to remember.

When one discharges an assumption, one sometimes says one *removes* them. One can remove as many instances of an assumption as one wants – from zero to several billions. In practice, one almost always wants to discharge as many as possible. In the following example there are two instances which are discharged:

5.2.2 Example. Derive $(P_1 \wedge P_2) \to (P_2 \wedge P_1)$.

Solution.

$$\dfrac{\dfrac{\dfrac{[P_1 \wedge P_2]}{P_2} \qquad \dfrac{[P_1 \wedge P_2]}{P_1}}{P_2 \wedge P_1}}{(P_1 \wedge P_2) \to (P_2 \wedge P_1)}$$

□

In the following example, which is also correct, there are no instances of the assumption which are discharged:

5.2.3 Example. Derive $P_1 \to \top$.

Solution.

$$\dfrac{\overline{\top}}{P_1 \to \top}$$

□

5.2.4 Exercise. Derive $(P_1 \wedge (P_2 \wedge P_3)) \to ((P_1 \wedge P_2) \wedge P_3)$.

For the implication elimination we have the following rule. It says that it is correct to conclude that n^2 is odd if one knows both that it is true that "if n is odd, then n^2 is odd" and that "n is odd".

$$\dfrac{\varphi \to \psi \qquad \varphi}{\psi} \tag{5.2.5}$$

5.2.6 Example. Derive $(\varphi \wedge \psi) \to \sigma$ from $\varphi \to (\psi \to \sigma)$.

Solution.

$$\dfrac{\dfrac{\varphi \to (\psi \to \sigma) \qquad \dfrac{[\varphi \wedge \psi]}{\varphi}}{\psi \to \sigma} \qquad \dfrac{[\varphi \wedge \psi]}{\psi}}{\dfrac{\sigma}{(\varphi \wedge \psi) \to \sigma}}$$

□

As you can see, it starts getting difficult to see which rule is applied where. One therefore puts some small markings to the right of the line. One writes I with the introduction rule and E with the elimination rule. Furthermore, one writes where the assumption is discharged by enumerating them. The previous tree is therefore given the following markings:

$$\dfrac{\dfrac{\dfrac{\varphi \to (\psi \to \sigma) \qquad \dfrac{[\varphi \wedge \psi]^1}{\varphi}\,{\scriptstyle \wedge E}}{\psi \to \sigma}\,{\scriptstyle \to E} \qquad \dfrac{[\varphi \wedge \psi]^1}{\psi}\,{\scriptstyle \wedge E}}{\sigma}\,{\scriptstyle \to E}}{(\varphi \wedge \psi) \to \sigma}\,{\scriptstyle \to I_1}$$

5.2.7 Exercise. Derive $\varphi \to (\psi \to \sigma)$ from $(\varphi \wedge \psi) \to \sigma$ and give the tree the correct markings.

5.2.8 Example. Derive $P_1 \to P_1$.

Solution. One could do it in the following way:

$$\dfrac{\dfrac{\dfrac{[P_1]^1 \qquad [P_1]^1}{P_1 \wedge P_1}\,{\scriptstyle \wedge I}}{P_1}\,{\scriptstyle \wedge E}}{P_1 \to P_1}\,{\scriptstyle \to I_1}$$

This seems to be unnecessarily long, though. Instead, one could do it like this if only one accepts that the "tree" which is symbolized by the vertical lines in the implication-introduction rules could be a single formula: P_1. It then works as both an assumption which is discharged and a formula which is derived under the assumption.

$$\dfrac{[P_1]^1}{P_1 \to P_1}\,{\scriptstyle \to I_1}$$

□

We will consider such derivations to be correct.

5.3 Disjunction

The derivation rules for disjunction will in general be dual to those of conjunction. For an n-ary disjunction we get:

The conjuncts are formulas which are glued together by applying $\vee$ to get more complex formulas.

1. n *introduction rules* which consist of the disjunction itself under the line, and above the line, in the i-th introduction rule, we have the i-th disjunct.

2. One *elimination rule*.

The introduction rules are thus like the conjunction elimination rules, but turned upside down.

$$\frac{\varphi_i}{\varphi_1 \vee \cdots \vee \varphi_n}\,. \tag{5.3.1}$$

It is tempting to do something similar with the elimination rule. Unfortunately, this would not work very well, since it would mean that we would get more formulas *under* the line, which would not work quite well with our other rules. Instead, the elimination rules should be like:

$$\frac{\varphi_1 \vee \cdots \vee \varphi_n \qquad \begin{array}{c}[\varphi_1]\\ \vdots\\ \sigma\end{array} \quad \cdots \quad \begin{array}{c}[\varphi_n]\\ \vdots\\ \sigma\end{array}}{\sigma} \tag{5.3.2}$$

The assumption φ_i is only allowed to be discharged in the subtree symbolized by

$$\begin{array}{c}[\varphi_i]\\ \vdots\\ \sigma.\end{array}$$

At first glance, this rule may seem somewhat difficult to read. The idea behind the rule is that if we know that $\varphi_1 \vee \cdots \vee \varphi_n$ is true, and we have derivations of σ from each and every one of $\varphi_1, \ldots, \varphi_n$, then σ must be true.

The vertical dots symbolize derivations precisely as in the case of implication introduction. The disjunction elimination also discharges assumptions. It is important to understand that discharging φ_i must *only* occur in the corresponding subtree; that is, the one which is symbolized by the dots under φ_i. In this subtree, however, one can discharge as many instances of φ_i as one wants – from none to millions.

For a binary disjunction we get the introduction rules:

$$\frac{\varphi}{\varphi \vee \psi} \qquad\qquad \frac{\psi}{\varphi \vee \psi} \tag{5.3.3}$$

and the elimination rule:

$$\frac{\varphi \vee \psi \qquad \begin{array}{c}[\varphi]\\ \vdots\\ \sigma\end{array} \qquad \begin{array}{c}[\psi]\\ \vdots\\ \sigma\end{array}}{\sigma} \tag{5.3.4}$$

For our nullary disjunction $\perp$ we get, according to the same pattern, 0 introduction rules, and the elimination rule:

$$\frac{\perp}{\sigma}\,. \tag{5.3.5}$$

The fact that we do not get any introduction rule should be interpreted by the fact that we are never allowed to conclude that $\perp$ symbolizes a true formula. The elimination rule means that if we have concluded that $\perp$ is true, then we can also conclude that σ is true, for any formula σ. We can look at this as saying that one might as well "give up" and interpret everything as true if one has succumbed to interpret $\perp$ as true. A better explanation I personally think is reasonable is to simply look at $\perp$ as a nullary disjunction and observe that the rule follow the pattern. It is *us* who decide what $\perp$ should mean, so we are free to say it should be a nullary disjunction, from which the rule follows.

5.3.6 Exercise. Show that $\varphi \vdash \varphi \vee \perp$ holds for any formula φ, i.e. that one can derive $\varphi \vee \perp$ from φ.

5.3.7 Exercise. Show that one can derive $\psi \vee \varphi$ from $\varphi \vee \psi$, i.e., that $\varphi \vee \psi \vdash \psi \vee \varphi$ holds for any formulas φ, ψ.

5.3.8 Exercise. Show that $\varphi \vee \perp \vdash \varphi$ is true for any formula φ, i.e., that one can derive φ from $\varphi \vee \perp$.

5.3.9 Exercise. Derive $\varphi \vee \varphi \to \varphi$. (In other words, construct a derivation without any undischarged assumption.)

5.3.10 Exercise. Derive $(\varphi \vee \psi) \vee \sigma \to \varphi \vee (\psi \vee \sigma)$.

$$\frac{}{\top}\,{\top}I \qquad\qquad \frac{\bot}{\sigma}\,{\bot}E$$

$$\frac{\varphi \qquad \psi}{\varphi \wedge \psi}\,{\wedge}I \qquad\qquad \frac{\varphi \wedge \psi}{\varphi}\,{\wedge}E \qquad \frac{\varphi \wedge \psi}{\psi}\,{\wedge}E$$

$$\frac{\varphi}{\varphi \vee \psi}\,{\vee}I \qquad \frac{\psi}{\varphi \vee \psi}\,{\vee}I \qquad\qquad \frac{\varphi \vee \psi \qquad \begin{array}{c}[\varphi]\\ \vdots \\ \sigma\end{array} \qquad \begin{array}{c}[\psi]\\ \vdots \\ \sigma\end{array}}{\sigma}\,{\vee}E$$

$$\frac{\begin{array}{c}[\varphi]\\ \vdots \\ \psi\end{array}}{\varphi \to \psi}\,{\to}I \qquad\qquad \frac{\varphi \to \psi \qquad \varphi}{\psi}\,{\to}E$$

$$\frac{\begin{array}{c}[\neg\sigma]\\ \vdots \\ \bot\end{array}}{\sigma}\,\text{RAA}$$

Figure 5.1: Derivation rules for natural deduction in propositional logic

5.4 Negation and equivalence

We do not formulate any rules for $\neg$ and $\leftrightarrow$, because these operations are defined in terms of others (Definition 4.1.3). The following derivation is therefore correct:

$$\dfrac{\dfrac{\dfrac{[\bot]^1}{\neg\bot}\,{\to}I_1}{\top\to\neg\bot}\,{\to}I_2 \qquad \dfrac{\dfrac{}{\top}\,\top I}{\neg\bot\to\top}\,{\to}I_3}{\top\leftrightarrow\neg\bot}\,\wedge I \tag{5.4.1}$$

To see this, you can substitute $\neg\bot$ by $\bot\to\bot$, and so on.

The rules we have seen so far are not enough if one would like, for example, to be able to derive the formula $P_1 \vee \neg P_1$ (you will be able to prove this in Exercise 7.3.9). Therefore, one adds a specific rule for such purpose, called RAA.

RAA = *reductio ad absurdum*. This could be a distortion of the phrase *deductio ad absurdum* (derivation of the impossible) which one finds in older texts.

$$\begin{array}{c}[\neg\sigma]\\ \vdots\\ \dfrac{\bot}{\sigma}\,\text{RAA}\end{array} \tag{5.4.2}$$

It deviates in its form from the previous rule. It is in fact a strengthening of the rule $\bot E$, since both allow us to conclude σ from $\bot$, but RAA allows us also to discharge as many instances of the assumption $\neg\sigma$ as we like. Using RAA one can derive $\varphi\vee\neg\varphi$ in the following way, for every formula φ:

5.4.3 Example. Derive $\varphi\vee\neg\varphi$.

Solution.

$$\dfrac{\dfrac{[\neg(\varphi\vee\neg\varphi)]^2 \qquad \dfrac{\dfrac{\dfrac{[\neg(\varphi\vee\neg\varphi)]^2 \qquad \dfrac{[\varphi]^1}{\varphi\vee\neg\varphi}\,\vee I}{\bot}\,{\to}E}{\neg\varphi}\,{\to}I_1}{\varphi\vee\neg\varphi}\,\vee I}{\bot}\,{\to}E}{\varphi\vee\neg\varphi}\,\text{RAA}_2$$

□

5.4.4 Exercise. What formula does one derive if one changes the last rule to $\to I$ above?

5.5 The formal point of view

We have now gone through all the rules there are and that are allowed in natural deduction. When correcting exams, though, it is apparent that many students invent their own additional rules. This is not allowed. We will prove theorems by natural deduction and the proofs of these theorems will use in an essential way that no other rules other than the ones we have collected in Figure 5.1 occur. You should learn them by heart, which means that you should both memorize them and understand how they are used. You should always mark every rule you use by its name, it makes it clearer both to yourself and the ones who read your derivation that the rules you use actually exist and are applicable. You can look at derivations as a sort of game. The point is that you succeed with the exercise by following the rules.

5.5.1 Definition. 1. By $\varphi_1,\ldots,\varphi_n \vdash \varphi$ we mean that there exists a derivation that concludes φ, according to the rules in Figure 5.1, where there are no undischarged assumptions except, possibly, $\varphi_1,\ldots,\varphi_n$.

It is not mandatory to make use of the formulas to the left of $\vdash$.

2. One says that such a derivation is "a derivation of φ from $\varphi_1, \dots, \varphi_n$".
3. When constructing such derivations, one says that one is "deriving φ from $\varphi_1, \dots, \varphi_n$".
4. By $\vdash \varphi$ it is meant, in particular, that there is a derivation concluding φ without any undischarged assumptions.
5. One says that such a thing is "a derivation of φ".
6. When constructing such a derivation, one says that one is "deriving φ".

Though we take such a formal point of view on derivations, the rules we have chosen are of course not randomly chosen. We have motivated the introduction of the rules which we have collected in Figure 5.1. In principle we could add more rules, but we are satisfied with those we have, since they are enough for what we are going to do. We will prove this in Chapter 8: the rules we have introduced is a *complete* system in the sense that everything which is *true in all interpretations*, and which can be expressed in the language we are studying, can also be *derived* through the rules we collected in Figure 5.1 (Completeness theorem 8.2.3, page 62).

5.6 Miscellaneous exercises

5.6.1 Exercise (from the exam on 2004-01-08)**.**
Give a complete derivation in natural deduction of the following formula:

$$(\neg\varphi \to \psi) \leftrightarrow (\varphi \vee \psi)$$

5.6.2 Exercise (from the exam on 2002-10-21)**.**
Give a complete derivation in natural deduction of the following formula:

$$\neg(P_1 \to P_2) \leftrightarrow P_1 \wedge \neg P_2$$

5.6.3 Exercise (from the exam on 2002-08-20)**.**
Give a complete derivation in natural deduction of the following formula:

$$(P_3 \to (P_1 \to P_2)) \leftrightarrow (P_3 \wedge P_1 \to P_2)$$

5.6.4 Exercise (from the exam on 2004-10-18)**.**
Give a complete derivation in natural deduction of the following formula:

$$((\varphi \to \psi) \to \psi) \leftrightarrow (\varphi \vee \psi)$$

5.6.5 Exercise (from the exam on 2005-01-07)**.**
Give a complete derivation in natural deduction of the following formula:

$$((\varphi \vee \psi) \wedge \neg\psi) \leftrightarrow (\varphi \wedge \neg\psi)$$

5.6.6 Exercise (from the exam on 2005-01-07)**.**

a) Find all mistakes in the following derivation. Specify them carefully!

$$\dfrac{\dfrac{\dfrac{[\varphi \vee \psi]^1}{\varphi}\,{\scriptstyle \vee E} \qquad \dfrac{[\varphi \vee \psi]^1}{\psi}\,{\scriptstyle \vee E}}{\varphi \wedge \psi}\,{\scriptstyle \wedge I}}{(\varphi \vee \psi) \to (\varphi \wedge \psi)}\,{\scriptstyle \to I_1}$$

b) Give examples of formulas φ, ψ such that there is a correct derivation of $(\varphi \vee \psi) \to (\varphi \wedge \psi)$. Motivate them carefully!

c) Show that if φ, ψ are formulas such that $\vdash (\varphi \vee \psi) \to (\varphi \wedge \psi)$ holds, then we also have $\varphi \vdash \psi$.

5.7 Summary

You have learnt what a *derivation in natural deduction* is. The difference between *discharged* and *undischarged* assumptions has been explained and you have learnt when an assumption may be discharged. The most important thing to remember for the rest of this course is the ability to construct a derivation, to decide whether a derivation is correct and the insight that only the given rules are allowed in such derivations.

Chapter 6

Soundness & Review exercises

6.1 Soundness

We have encountered the expressions

$$\varphi_1, \ldots, \varphi_n \vDash \varphi\,, \tag{6.1.1}$$

$$\varphi_1, \ldots, \varphi_n \vdash \varphi\,. \tag{6.1.2}$$

Although they look much alike, they are two completely different things. The former (6.1.1) means that the formula φ is true in certain interpretation, while the latter (6.1.2) means that φ can be derived according to certain rules. The reason the notation is so similar is that (6.1.1) and (6.1.2) are in fact equivalent. This says that, even though they mean different things, they always happen to be true at the same time. Here we will show that (6.1.2) implies (6.1.1). In Chapter 8 we will show that the converse is also true.

What a *derivation* is has been defined inductively in Chapter 5, even if we have not written the inductive definition properly. Now it is the time to do that, since we will prove propositions about *all derivations* and we would need to do it through an inductive argument.

A formula is a derivation, namely, the derivation of the formula from itself. The rules in Figure 5.1 construct the rest of the derivations. If, for example, $\mathcal{D}'$ and $\mathcal{D}''$ are derivations whose conclusions are φ_1 respectively φ_2, then

$$\frac{\mathcal{D}' \qquad \mathcal{D}''}{\varphi_1 \wedge \varphi_2}\,{\scriptstyle \wedge I} \tag{6.1.3}$$

is a derivation. Another example is the following: if $\mathcal{D}'$ is a derivation whose conclusion is ψ, then

$$\frac{\mathcal{D}'}{\varphi \to \psi}\,{\scriptstyle \to I} \tag{6.1.4}$$

– possibly with one or more assumptions of φ marked as discharged – is also a derivation.

The two examples we have seen above should be enough to understand how one defines the set of derivations inductively. This definition automatically gives principles for doing proofs by induction on the structure of derivations. We will do such a proof now that we reach one of the most important theorems of this course.

6.1.5 Theorem (soundness theorem)**.** *Consider a derivation in natural deduction. Then the concluded formula is true in all interpretations in which the undischarged assumptions hold.*

Proof. Let us show this by induction on the structure of derivations. This means that we will assume, before the proof:

1. that $\mathcal{D}$ is a derivation in natural deduction concluding φ,
2. (inductive hypothesis) that the theorem is true for all derivations we have encountered in our construction of $\mathcal{D}$.

Most of the cases have similar proofs, so you can, if you want you, be satisfied by reading the cases 2, 3, 7 and 9.

Our task is to show that $[\![\varphi]\!]^{\mathcal{A}} = 1$ for all interpretations $\mathcal{A}$ in which the undischarged assumptions are true. We go through various cases depending on which rule is the last rule that has been applied in $\mathcal{D}$.

Case 0: $\mathcal{D}$ is of the form

$$\varphi \tag{6.1.6}$$

Then φ is both the conclusion and the undischarged assumption, so in this case the claim is obvious.

Case 1: $\mathcal{D}$ is of the form

$$\frac{}{\top}\,\top I \tag{6.1.7}$$

Then $\varphi = \top$ and we have $[\![\varphi]\!] = [\![\top]\!] = 1$.

Case 2: $\mathcal{D}$ is of the form

$$\frac{\mathcal{D}'}{\varphi}\,\bot E \tag{6.1.8}$$

where $\mathcal{D}'$ is a derivation concluding $\bot$. The inductive hypothesis gives that the theorem is true for it. That is, $\bot$ is true in all interpretations in which all undischarged assumptions are true. But $\bot$ is not true in any interpretation, so there cannot be any interpretation in which all undischarged assumptions are true. Therefore, it holds that φ is true in all interpretations in which all undischarged assumptions are true – since there are zero such interpretations.

One says that something is *vacuously* true when it is true *for all* elements with a certain property because there are no elements with that property.

Case 3: $\mathcal{D}$ is of the form

$$\frac{\mathcal{D}' \qquad \mathcal{D}''}{\varphi_1 \wedge \varphi_2}\,\wedge I \tag{6.1.9}$$

where $\mathcal{D}'$ concludes φ_1 and $\mathcal{D}''$ concludes φ_2. Every $\mathcal{A}$ which interprets all undischarged assumptions in $\mathcal{D}$ as true interprets also the undischarged assumptions in $\mathcal{D}'$ and $\mathcal{D}''$ as true. Thus, $[\![\varphi_1]\!] = [\![\varphi_2]\!] = 1$ for all such interpretations. But then $[\![\varphi]\!] = [\![\varphi_1 \wedge \varphi_2]\!] = [\![\varphi_1]\!] \wedge [\![\varphi_2]\!] = 1 \wedge 1 = 1$.

Case 4: $\mathcal{D}$ is of the form

$$\frac{\mathcal{D}'}{\varphi}\,\wedge E \tag{6.1.10}$$

where the conclusion of $\mathcal{D}'$ is $\varphi_1 \wedge \varphi_2$. Thus, we have $\varphi = \varphi_i$ for $i = 1$ or $i = 2$. Every $\mathcal{A}$ which interprets all undischarged assumptions in $\mathcal{D}$ as true also interprets the undischarged assumptions in $\mathcal{D}'$ as true. Thus $[\![\varphi_1 \wedge \varphi_2]\!] = 1$ for such interpretations. But then $[\![\varphi_1]\!] = [\![\varphi_2]\!] = 1$, so $[\![\varphi]\!] = 1$.

Case 5: $\mathcal{D}$ is of the form

$$\frac{\mathcal{D}'}{\varphi_1 \vee \varphi_2}\,\vee I \tag{6.1.11}$$

where the conclusion of $\mathcal{D}'$ is φ_i for $i = 1$ or $i = 2$. Every $\mathcal{A}$ which interprets all undischarged assumptions in $\mathcal{D}$ as true also interprets the undischarged assumptions in $\mathcal{D}'$ as true. Thus, $[\![\varphi_i]\!] = 1$ for such $\mathcal{A}$. It follows that $[\![\varphi]\!] = [\![\varphi_1 \vee \varphi_2]\!] = [\![\varphi_1]\!] \vee [\![\varphi_2]\!] \geq [\![\varphi_i]\!] = 1$.

Case 6: $\mathcal{D}$ is of the form

$$\frac{\mathcal{D}' \qquad \mathcal{D}_1 \qquad \mathcal{D}_2}{\varphi}\,\vee E \tag{6.1.12}$$

where

- $\mathcal{D}'$ concludes $\varphi_1 \vee \varphi_2$,
- $\mathcal{D}_1$ and $\mathcal{D}_2$ concludes φ,
- (possibly) some undischarged assumptions of φ_1 in $\mathcal{D}_1$ have been marked as discharged and

- (possibly) some undischarged assumption of φ_2 in $\mathcal{D}_2$ have been marked as discharged.

The inductive hypothesis says that the theorem is true for $\mathcal{D}', \mathcal{D}_1, \mathcal{D}_2$.

Consider now an interpretation $\mathcal{A}$ in which all undischarged assumptions in $\mathcal{D}$ are true. Then all undischarged assumptions in $\mathcal{D}'$ are also true, so $[\![\varphi_1 \vee \varphi_2]\!] = 1$. Since $[\![\varphi_1 \vee \varphi_2]\!] = [\![\varphi_1]\!] \vee [\![\varphi_2]\!]$, one of the disjuncts has to be 1, say $[\![\varphi_i]\!] = 1$. Consider now $\mathcal{D}_i$. The undischarged assumptions in this are either φ_i, which is true in $\mathcal{A}$, or they are also undischarged in $\mathcal{D}$. Thus, all undischarged assumptions in $\mathcal{D}_i$ are true in $\mathcal{A}$. It follows that the conclusion of $\mathcal{D}_i$ is true in $\mathcal{A}$, but this is precisely φ.

Case 7: $\mathcal{D}$ is of the form

$$\frac{\mathcal{D}'}{\varphi_1 \to \varphi_2} \to I \tag{6.1.13}$$

where $\mathcal{D}'$ concludes φ_2 and (possibly) some assumptions of φ_1 in $\mathcal{D}'$ have been marked as discharged. Consider now an interpretation in which all undischarged assumptions in $\mathcal{D}$ are true. We will show that $[\![\varphi_1 \to \varphi_2]\!] = 1$; that is to say, if $[\![\varphi_1]\!] = 1$ then $[\![\varphi_2]\!] = 1$. This follows from the fact that, if $[\![\varphi_1]\!] = 1$, then all undischarged assumptions in $\mathcal{D}'$ are true, and therefore also φ_2 is true according to the inductive hypothesis.

Case 8: $\mathcal{D}$ is of the form

$$\frac{\mathcal{D}' \qquad \mathcal{D}''}{\varphi} \to E \tag{6.1.14}$$

where $\mathcal{D}'$ is a derivation of $\psi \to \varphi$ and $\mathcal{D}''$ is a derivation of ψ. Every $\mathcal{A}$ which interprets the undischarged assumptions in $\mathcal{D}$ as true also interprets the undischarged assumptions in $\mathcal{D}'$ and $\mathcal{D}''$ as true. Thus, $\psi \to \varphi$ and ψ are true in such interpretation. It follows that φ is true in that interpretation.

Case 9: $\mathcal{D}$ is of the form

$$\frac{\mathcal{D}'}{\varphi} \text{RAA} \tag{6.1.15}$$

where $\mathcal{D}'$ concludes $\bot$ and (possibly) some assumptions of $\neg\varphi$ in $\mathcal{D}'$ are marked as discharged. Then $\bot$ is true in all interpretations in which all asumptions in $\mathcal{D}$ are true at the same time as $\neg\varphi$ is true. But $\bot$ is not true in any interpretation, so it follows that $\neg\varphi$ is not true in any interpretation in which all undischarged assumptions in $\mathcal{D}$ are true. Therefore, φ must be true in all such interpretations. □

One can formulate the soundness theorem in another way, which at first glance might seem stronger. To do this we need some definitions. We will generalize $\vDash$ and $\vdash$ so that we allow not only finitely many formulas in the left, but even infinitely many.

▶ **6.1.16 Definition.** If $\Gamma \subseteq \text{Form}$, then $\Gamma \vDash \varphi$ means that every model of Γ is a model of φ.

Definition 4.2.39 defines what a model is.

▶ **6.1.17 Definition.** If $\Gamma \subseteq \text{Form}$, then $\Gamma \vdash \varphi$ means that φ can be derived without any other rules than those given in Figure 5.1 and without any other undischarged assumptions, except, possibly, formulas in Γ. $\Gamma \nvdash \varphi$ means that no such derivation exists.

6.1.18 Exercise. How can one express $\{\varphi_1, \ldots, \varphi_n\} \vDash \varphi$ and $\{\varphi_1, \ldots, \varphi_n\} \vdash \varphi$ using the old notation?

6.1.19 Theorem (soundness theorem in another formulation)**.** $\Gamma \vdash \varphi \;\Rightarrow\; \Gamma \models \varphi$

Proof. Assume that $\Gamma \vdash \varphi$; that is, there is a derivation $\mathcal{D}$ of φ where the undischarged assumptions $\gamma_1, \ldots, \gamma_n$ are all in Γ. For all interpretations $\mathcal{A}$ in which all formulas in Γ are true, we have that $\gamma_1, \ldots, \gamma_n$ are true, and thus, from the soundness theorem 6.1.5 it follows that in all such interpretations φ is also true, which was what we had to show. □

Special case of the soundness theorem: $\vdash \varphi \;\Rightarrow \models \varphi$ says that only tautologies can be derived without undischarged assumptions.

With the help of the soundness theorem we can sometimes easily show results about interpretations.

6.1.20 Example. Show that if $P_1 \wedge P_2$ is true in $\mathcal{A}$, then $P_2 \wedge P_1$ is true in $\mathcal{A}$.

Solution. Follows from soundness and (5.1.8). □

6.1.21 Example. Show that for all formulas φ, ψ, we have that $((\varphi \vee \psi) \wedge \neg\psi) \leftrightarrow (\varphi \wedge \neg\psi)$ is a tautology.

Solution. It follows from the soundness theorem and the answer to Exercise 5.6.5. □

6.1.22 Exercise. Show that $P_1 \to ((P_2 \vee P_3) \wedge (P_4 \vee P_5) \to P_1)$ is a tautology.

6.1.23 Example. Show that if $\vdash \varphi \leftrightarrow \psi$ then $\varphi \approx \psi$.

Solution. Assume that $\vdash \varphi \leftrightarrow \psi$. According to the soundness theorem, we have $\models \varphi \leftrightarrow \psi$. The rest follows from Exercise 4.2.31. □

One can also show that it is impossible to derive certain formulas by using only the rules in Figure 5.1.

6.1.24 Example. Show that $\nvdash (P_1 \vee P_2)$; that is, it is not possible to derive the formula using the rules in Figure 5.1.

Solution. Assume that one could derive $P_1 \vee P_2$. Then it should, according to the soundness theorem, be a tautology. But it is not, since it is false in the interpretation in which P_1 and P_2 are both intepreted as false propositions. □

6.1.25 Exercise. Show that one cannot derive $(P_1 \vee P_2) \to P_1$.

▶ **6.1.26 Definition.** To say that Γ is inconsistent means that $\Gamma \vdash \bot$. By Γ is consistent it is meant that $\Gamma \nvdash \bot$.

6.1.27 Example. Show that $\{P_1, P_2, P_3, P_4\}$ is consistent.

Solution. Assume that $\{P_1, P_2, P_3, P_4\}$ was inconsistent; that is,

$$P_1, P_2, P_3, P_4 \vdash \bot\,.$$

Then, according to the soundness theorem, $P_1, P_2, P_3, P_4 \models \bot$. But if $\mathcal{A}$ interprets P_1, P_2, P_3, P_4 as true, then we still have $[\![\bot]\!]^{\mathcal{A}} = 0$, which contradicts $P_1, P_2, P_3, P_4 \models \bot$. □

6.1.28 Exercise (from the exam on 2002-10-21)**.** Decide if the following subsets of Form are consistent.

a) $\{P_1 \vee P_2, P_2 \vee \neg P_3, \neg P_3 \vee \neg P_4, P_3 \vee \neg P_1, \neg P_2 \vee P_4\}$

b) $\{P_1 \to P_2, P_2 \to P_3, P_3 \to \neg P_1, P_4 \to P_2, P_3 \to \neg P_4, \neg P_4 \to P_1\}$

6.1.29 Exercise (from the exam on 2005-01-07)**.** Give examples of formulas φ, ψ for which no correct derivation of $(\varphi \vee \psi) \to (\varphi \wedge \psi)$ exists. Cf. Exercise 5.6.6.

6.1.30 Example. Show that no derivation of $P_1 \vee \neg P_1$ can be concluded with an introduction rule.

Solution. The only introduction rule which could possibly have been used is the or-introduction rule. Assume then that we have a derivation of $P_1 \vee \neg P_1$ concluded by an or-introduction rule. If the last step in the derivation is removed, we would have a derivation of either P_1 or $\neg P_1$. In the first case, we should have, according to the soundness theorem, $\models P_1$, which we do not have, since P_1 can be interpreted as false. In the other case, we would have $\models \neg P_1$, which we do not have, since P_1 can be interpreted as true, in which case $[\![\neg P_1]\!] = 0$. □

6.1.31 Exercise. Show that no derivation of $P_1 \vee (P_2 \vee P_3)$ from $(P_1 \vee P_2) \vee P_3$ can conclude with an introduction rule. (It follows that Exercise 5.3.10 cannot be solved with a tree that concludes in two introduction rules.)

6.1.32 Exercise. Show that a derivation of $\varphi \vee (\psi \vee \sigma)$ from $(\varphi \vee \psi) \vee \sigma$ can conclude with an introduction rule *for some* choices of φ, ψ, σ, but *not* for all.

6.1.33 Exercise. Show that no derivation without undischarged assumptions can end with $\bot$-elimination.

6.1.34 Exercise. Show that one cannot solve Exercise 5.3.7 with a tree ending in a introduction rule.

6.1.35 Exercise. Show that one cannot derive $(P_1 \vee P_2) \to (P_1 \vee P_2)$ using a tree ending in two introduction rules, but it is possible to derive it in another way.

6.1.36 Exercise. Show that if one derives $P_1 \to P_1$ by a tree ending in an introduction rule, one must discharge at least one assumption.

6.1.37 Exercise.

a) Show that $\varphi_1, \ldots, \varphi_n \models \varphi \iff [\![\varphi_1]\!] \wedge \cdots \wedge [\![\varphi_n]\!] \leq [\![\varphi]\!]$ for all interpretations.

b) Show that $\varphi_1, \ldots, \varphi_n, \varphi \models \psi \iff \varphi_1, \ldots, \varphi_n \models \varphi \to \psi$ follows from the Galois connection (2.2.3). This can thus be used as an alternative proof for soundness in the case of implication introduction.

6.1.38 Exercise. Complete by yourself the cases $\wedge I$ and $\to I$ of the proof of the soundness theorem without looking in the book.

6.2 Summary

The only result of this chapter is the *soundness theorem.* It is one of the most important theorems in propositional logic. What is important to take with you for the rest of the course is the ability to use the soundness theorem to see that certain attempts of deriving a formula are doomed to fail (this way you will find the viable paths more easily), as well as the ability to see that certain formulas cannot be derived at all if the formulas in a given set should only to be used as undischarged assumptions. A special case of this is the proof of consistency: you should be able to prove using the soundness theorem that a given set of formulas is consistent. Another special case is the fact that only tautologies can be derived without undischarged assumptions.

If you are still in doubt about whether some of the rules in natural deduction are correct, take another look at the proof of the soundness theorem. There it is in fact proven that the rules are correct in a certain sense – namely, that they are in agreement with the semantics.

6.3 Review exercises

6.3.1 Exercise. Show that in any Boolean algebra we have $a \to b = 1$ if and only if $a \leq b$.

Hint. Use (2.2.3).

6.3.2 Exercise. Describe what this function does (defined on natural numbers):

$$f(a, 0) \stackrel{\text{def}}{=} a$$
$$f(a, s(n)) \stackrel{\text{def}}{=} s(f(a, n))$$

One says that φ is *atomic* if $a(\varphi) = 0$.

6.3.3 Exercise. We define the *number of operations* in a formula in the following way, as a function from Form to $\mathbb{N}$:

$$a(P_j) = 0$$
$$a(\top) = 1$$
$$a(\bot) = 1$$
$$a(\varphi \wedge \psi) = s(a(\varphi) + a(\psi))$$
$$a(\varphi \vee \psi) = s(a(\varphi) + a(\psi))$$
$$a(\varphi \to \psi) = s(a(\varphi) + a(\psi))$$

Compute $a(\top \leftrightarrow \neg P_1)$. You do not need to show every step.

6.3.4 Exercise. Are these propositions about formulas in Form true?

a) $(\bot \wedge \bot) = \bot$

b) $(P_1 \leftrightarrow \bot) = (\neg P_1 \wedge (\bot \to P_1))$

c) $(\bot \wedge \bot) \approx \bot$

d) $(P_1 \leftrightarrow \bot) \approx (\neg P_1 \wedge (\bot \to P_1))$

6.3.5 Exercise (from the exam on 2004-01-08)**.** Give a formula φ which solves the following problem:

- if P_1 is interpreted as false, φ is also interpreted as false,
- if P_1 is interpreted as true, φ is interpreted as true if and only if *precisely one* of P_2 and P_3 are interpreted as true.

6.3.6 Exercise (from the exam on 2005-08-23)**.**
Give a complete derivation in natural deduction of the following formula:

$$((\varphi \vee \psi) \to \sigma) \leftrightarrow ((\varphi \to \sigma) \wedge (\psi \to \sigma))$$

You have seen this earlier in the text, but try not to look at it.

6.3.7 Exercise (from the exam on 2004-08-17)**.**
Give a complete derivation in natural deduction of the following formula:

$$\neg\neg(\varphi \vee \neg\varphi)$$

6.3.8 Exercise (from the exam on 2003-10-20)**.**
Give a complete derivation in natural deduction of the following formula:

$$\varphi \wedge (\psi \vee \sigma) \leftrightarrow (\varphi \wedge \psi) \vee (\varphi \wedge \sigma)$$

6.3.9 Exercise (from the exam on 2003-08-19)**.**
Give a complete derivation in natural deduction of the following formula:

$$((P_3 \to P_1) \to P_2) \leftrightarrow ((\neg P_3 \to P_2) \wedge (P_1 \to P_2))$$

6.3.10 Exercise (from the exam on 2003-01-09)**.**
Give a complete derivation in natural deduction of the following formula, without using RAA:

$$(\neg P_1 \to \neg P_2) \leftrightarrow \neg(\neg P_1 \wedge P_2)$$

Chapter 7

Normal deductions

The purpose of this chapter is to give you the chance to polish your ability to construct derivations and learn how to look for them more systematically. The theory is rather extensive, but several proofs have been put in an appendix (Normalization proofs, p. 113), and you do not have to read them if you are not particularly interested. The important thing is that you understand how you can search more efficiently for derivations with knowledge about the so called normal derivations. With such knowledge, one can discover which attempts are dead ends and get good hints about which possibilities one should investigate.

7.1 Introduction

For a derivation to be *non-normal* we say, loosely speaking, that it contains detours (we shall soon make this precise). For example, a derivation of the following form is not normal:

$$\dfrac{\dfrac{\begin{matrix}[\varphi]\\ \vdots \\ \psi\end{matrix}}{\varphi \to \psi}\,{\scriptstyle \to I} \qquad \begin{matrix}\vdots\\ \varphi\end{matrix}}{\psi}\,{\scriptstyle \to E} \tag{7.1.1}$$

since one can avoid the introduction of $\to$ through the transformation

$$\begin{matrix}\vdots\\ \varphi\\ \vdots\\ \psi.\end{matrix} \tag{7.1.2}$$

In (7.1.1) one simply replaces each discharged assumption of φ by the sub-derivation which concludes φ and take away the last two steps in the derivation. One says that one *normalizes* the derivation when one straightens out such detours. The observation is that when we introduce a logical operation and then immediately remove it, it is unnecessary to have introduced it at all. We say that a derivation is normal if it does not contain such an unnecessary complication. We need to make more precise what one calls an "unnecesasry complication". To facilitate this, we introduce some new terminology.

▶ **7.1.3 Definition.** In every derivation rule (Figure 5.1, page 40) we call the formulas above the line *premises* and those which are underneath the line *conclusions*. In the elimination rule, we call those premises whose operations are eliminated *main premises* and the others are called *side premises* or *minor premises*.

7.1.4 Exercise. Go through all the derivation rules and mark where are the premises. Which of these are main premises, respectively side premises?

Hint. There are only three side premises in the table.

It is not really the formula *in itself* which is a premise, respectively, a conclusion, but it is rather its *place* which decides that. One can thus not say that the premises are a subset of Form and so on. Often the same formula is conclusion in one rule and premise in the next one (namely, when it is not an assumption nor the conclusion of the entire derivation).

The definition of *normal* we use here ressembles the one given by Seldin, but is somewhat simplified. It is easier to understand and sufficient for this course, but has some disadvantages in other applications compared to the usual definition.

7.1.5 Definition. A derivation (in propositional logic) is *normal* if no main premise in an elimination rule is the conclusion in any other rule but $\wedge E$ or $\to E$.

One can ask if there are many derivations which are normal according to this definition. In fact, every derivation can be "normalized", which means that one gradually transforms it until a normal derivation occurs, similarly to the case of normalizing Boolean expressions (Section 1.5). This is what Theorem 7.2.6 says. We shall see first how, step by step, one approaches normal derivations through succesive transformations.

7.2 Glivenko's theorem and normalization

We start by observing that one can always do without RAA, except, possibly, in the last step. We shall see that if RAA is used further up in the derivation, one can transform the derivation so that the usage of RAA is pushed downwards. By repeating this process, one gets a derivation where RAA is not used at all except, possibly, in the last step.

Consider, for example, the following derivation:

$$\dfrac{\dfrac{\dfrac{\neg\neg(\varphi\to\psi) \qquad [\neg(\varphi\to\psi)]^1}{\bot}\,{\to}E}{\varphi\to\psi}\,\mathrm{RAA}_1 \qquad \varphi}{\psi\,.}\,{\to}E \tag{7.2.1}$$

It is not normal, since the main premise in the last elimination rule is the conclusion of RAA. But it can be transformed to

$$\dfrac{\dfrac{\neg\neg(\varphi\to\psi) \qquad \dfrac{\dfrac{[\neg\psi]^1 \qquad \dfrac{[\varphi\to\psi]^2 \quad \varphi}{\psi}\,{\to}E}{\bot}\,{\to}E}{\neg(\varphi\to\psi)}\,{\to}I_2}{\bot}\,{\to}E}{\psi}\,\mathrm{RAA}_1 \tag{7.2.2}$$

which is indeed a bigger derivation, but where the usage of RAA has been pushed down to become the last step, making the resulting derivation normal. The fact that this can always be done is the content of the following theorem. The theorem in itself does not guarantee that the result becomes normal, only that the RAA can be pushed down. In a later theorem we shall also prove that one can always get a normal result.

7.2.3 Theorem (Glivenko's theorem)**.** *Every derivation can be transformed so that in the end it becomes a derivation in which RAA does not occur except possible at the last step, and in which all undischarged assumptions occurred (as undischarged assumptions) in the original derivation.*

Proof. See the Normalization proof in the appendix (p. 113). □

A corollary, which in itself is sufficient to prove Glivenko's theorem, is the following:

7.2.4 Theorem. *If $\Gamma \vdash \bot$, then there exists an RAA-free derivation from Γ to $\bot$.*

Proof. Assume that $\Gamma \vdash \bot$. According to Glivenko's theorem, there is a derivation from Γ to $\bot$ in which RAA does not occur except, possibly, at the last step. But if this is the case, the last step can be removed if possible assumptions of $[\neg\bot]$ are replaced by derivations:

$$\frac{[\bot]}{\neg\bot\,.}\to I \qquad (7.2.5)$$

□

RAA-free derivations are often called *intuitionistic* since the validity of the rule RAA is questioned by intuitionism. Intuitionism is a school of thought within mathematical philosophy which bears suspicions towards the way mathematicians handle infinity. The semantics we will go through in the next part (Predicate logic) is rejected by intuitionism, which chooses instead to explain the meaning of logical operations in another way. According to this explanation, one cannot motivate the fact that RAA is a correct rule, so it is excluded from intuitionistic logic. Intuitionistic logic has been shown later to have applications in other contexts as well.

A consequence of this theorem is that when one looks for derivations, one can consider the usage of RAA only at the end of the derivation, if one uses it at all. Above the last rule, the derivation will be RAA-free.

We now know that one can always transform derivations so that any possible use of RAA is pushed downwards until it is only used in the last step. In a similar fashion, one can go through the sort of transformations, examples of which have been given in (7.1.1)→(7.1.2), to reach a normal derivation. This is the content of the following theorem.

7.2.6 Theorem (weak normalization)**.** *Every derivation can be transformed in such a way that a normal derivation is reached, in which all undischarged assumptions existed already (as undischarged assumptions) in the original derivation. If the original derivation was RAA-free then the resulting derivation consists only on the rules that were used in the original one.*

Proof. See the Normalization proof in the appendix (p. 115). □

It follows immediately from the theorem that if $\Gamma \vdash \varphi$, there exists a normal derivation of φ in which all undischarged assumtions belong to Γ. That $\Gamma \vdash \varphi$ means exactly that there is a derivation of φ in which all undischarged assumptions are in Γ. According to the result above, one can normalize a derivation in a way that all undischarged assumptions already existed in the original one. The theorem can also be applied to answer questions of the type: can one derive $\neg(P_1 \wedge P_2) \to (\neg P_1 \vee \neg P_2)$ only by using the rules $\vee I$, $\to I$, $\to E$ and $\wedge E$? The answer is no, because if one could succeed in this, one could also normalize a derivation and get a normal derivation without undischarged assumptions in which only the rules $\vee I$, $\to I$, $\to E$ and $\wedge E$ are used. That no such normal derivation exists is something you will hopefully be able to prove yourself after reading some of the following section.

7.3 Applications

Now that we know that derivations can always be normalized, we know also that when looking for a derivation, it is sufficient to just look for normal derivations. This means that we can limit our search quite severely. The following theorems are used to get an overview of how normal derivations look like. The first says, for instance, that if we search for a normal derivation without undischarged assumptions, we will not be able to end with an elimination rule

Remember that $\bot$ is a subformula of, for example, $\neg P_1$.

7.3.1 Theorem. *If a normal derivation ends in an elimination rule, then the main premise is a subformula in some undischarged assumption.*

One can also formulate this proof as an induction proof over the structure of the derivation.

Proof. The main premise in the last elimination rule cannot be the conclusion of any other rule than $\wedge E$ or $\to E$ (this is required for the derivation to be called normal). This also holds for the main premise in the row above. The same holds for every row upwards in the derivation. When we follow the derivation upwards along the main premise we will thus only pass through the last rule and rules of the type $\wedge E$ and $\to E$. None of these rules discharge assumptions about their main premises, so at last we will reach a main premise which is an undischarged assumption. Every conclusion in the rules we have passed through is a subformula in the main premise of the same rule, and hence it follows that the conclusion of the derivation is a subformula of the undischarged assumption we have reached. □

7.3.2 Exercise. Show that if a normal derivation ends with $\bot E$, there has to be an undischarged assumption which has $\bot$ as a subformula.

7.3.3 Exercise. Show that if $\vdash \varphi$ there is a normal derivation of φ which concludes either with RAA or an introduction rule.

7.3.4 Exercise. Can one derive $\neg(P_1 \wedge P_2) \to (\neg P_1 \vee \neg P_2)$ solely by the use of the rules $\vee I$, $\to I$, $\to E$ and $\wedge E$?

The following theorem shows which formulas one has to use in a RAA-free derivation:

7.3.5 Theorem (subformula property)**.** *In every normal derivation without RAA, every formula is a subformula of either the conclusion or one of the undischarged assumptions of the derivation.*

Proof. Since every subderivation of a normal derivation is normal, we can prove the result by induction over the structure of normal derivations.

If the last rule is $\top I$ the result is obvious. If the last rule is $\bot E$ then the inductive hypothesis says that every formula, except the conclusion, is a subformula of the undischarged assumptions or of $\bot$. It is therefore sufficient to show that $\bot$ is a subformula of some undischarged assumption. But this was done in Exercise 7.3.2.

If the last rule is an introduction rule, which does not discharge any assumption, the result follows from the induction hypothesis and the fact that the premises in the last rule are subformulas of the conclusion.

In the case of $\to I$, whose conclusion is $\varphi \to \psi$, the result follows from what the inductive hypothesis says about the last formula, namely, that in fact every formula is a subformula of φ, ψ or of some assumption which is left undischarged by the last rule. Since both φ and ψ are subformulas of the conclusion of the derivation, the result follows.

In the case of $\wedge E$ and $\to E$ we can use the previous theorem to assert that the main premise is a subformula of an undischarged assumption. The side premises in $\to E$ are subformulas of the conclusion. The rest follows by inductive hypothesis.

Only the case of $\vee E$ is left. Assume therefore that the derivation looks like this:

$$\dfrac{\begin{matrix} \vdots \\ \varphi \vee \psi \end{matrix} \qquad \begin{matrix} [\varphi] \\ \vdots \\ \sigma \end{matrix} \qquad \begin{matrix} [\psi] \\ \vdots \\ \sigma \end{matrix}}{\sigma}\,\vee E \tag{7.3.6}$$

The inductive hypothesis says:

1. that every formula in the left is a subformula of one of the undischarged assumptions or of $\varphi \vee \psi$,

2. that every formula in the middle is a subformula of an undischarged assumption or of φ or σ,

3. that every formula in the right is a subformula of an undischarged assumption or of ψ or σ.

We can then draw the conclusion that every formula is a subformula of an undischarged assumption, or of $\varphi \vee \psi$ or of σ. It remains to show that $\varphi \vee \psi$ is a subformula of an undischarged assumption, but this follows from the previous theorem. □

7.3.7 Exercise. Show that one cannot possibly derive P_1 without any undischarged assumptions (this follows easily from the soundness theorem, but try to do it with the methods of this chapter).

7.3.8 Exercise. Show that one cannot possibly derive $\neg P_1$ without any undischarged assumptions (this exercise is also easily solved by the soundness theorem, but it is possible to do it with the methods of this chapter as well).

7.3.9 Exercise. Show that one cannot derive $P_1 \vee \neg P_1$ without using RAA.

Example 5.4.3 shows how it can be done *with* RAA.

7.3.10 Exercise. Can one derive $\neg(P_1 \wedge P_2) \to (\neg P_1 \vee \neg P_2)$ without RAA?

7.3.11 Exercise. Show that every RAA-free normal derivation of $(P_1 \vee P_2) \vee P_3$ from $P_1 \vee (P_2 \vee P_3)$ contains only the rules $\vee I$ and $\vee E$. More generally: show that if Γ contains only formulas which do not have any other logical connectives besides $\vee$, and φ is also such a formula and can be derived from Γ without RAA, then there is a normal derivation of φ from Γ which only uses $\vee I$ and $\vee E$.

Think how useful results of the type of Exercise 7.3.11 are! If one wants to search for a RAA-free derivation, one has, at any point, only two rules to choose between. This means that to construct such a derivation is almost automatic.

The following theorem shows that it is nearly always sufficient to search for derivations which end with an introduction rule.

7.3.12 Theorem. *Assume that* $\Gamma \vdash \varphi$.

1. *If φ is a $\wedge$-formula, then there is a normal derivation from Γ to φ which ends with $\wedge I$.*

2. *If φ is a $\to$-formula, then there is a normal derivation from Γ to φ which ends with $\to I$.*

Proof. Assume that $\varphi = \varphi_1 \wedge \varphi_2$ and that we have a derivation

$$\begin{array}{c} \vdots \\ \varphi_1 \wedge \varphi_2 \end{array} \tag{7.3.13}$$

whose undischarged assumptions are in Γ. Then the following is also such a derivation:

$$\dfrac{\dfrac{\begin{array}{c}\vdots\\ \varphi_1 \wedge \varphi_2\end{array}}{\varphi_1}\wedge E \qquad \dfrac{\begin{array}{c}\vdots\\ \varphi_1 \wedge \varphi_2\end{array}}{\varphi_2}\wedge E}{\varphi_1 \wedge \varphi_2\,.}\wedge I \tag{7.3.14}$$

We can now normalize the sub-derivations which conclude φ_1 respectively φ_2 and obtain, thereby, a normal derivation.

Assume instead that $\varphi = \varphi_1 \to \varphi_2$ and that we have a derivation

$$\begin{array}{c} \vdots \\ \varphi_1 \to \varphi_2 \end{array} \tag{7.3.15}$$

where the undischarged assumptions are in Γ. Then the following is also such a derivation:

$$\dfrac{\dfrac{\begin{array}{c}\vdots\\ \varphi_1 \to \varphi_2\end{array} \qquad [\varphi_1]}{\varphi_2}\to E}{\varphi_1 \to \varphi_2\,.}\to I \tag{7.3.16}$$

We can now normalize the sub-derivations which conclude φ_2 and thereby get a normal derivation. □

To understand this example you should take pen and paper and construct the derivation while you are reading.

7.3.17 Example. Look for a derivation of $(\varphi \vee \psi) \leftrightarrow \neg(\neg\varphi \wedge \neg\psi)$.

Solution. Since this is a $\wedge$-formula ($\leftrightarrow$ is defined as such) we know that it should end with a $\wedge I$. So what we are left with is to search for a derivation of

$$(\varphi \vee \psi) \to \neg(\neg\varphi \wedge \neg\psi) \tag{7.3.18}$$

and

$$\neg(\neg\varphi \wedge \neg\psi) \to (\varphi \vee \psi)\,. \tag{7.3.19}$$

Let us start with the first one. It is a $\to$-formula, so the derivation should conclude with $\to I$. We are left with searching for a derivation from $\varphi \vee \psi$ to $\neg(\neg\varphi \wedge \neg\psi)$. Again, the conclusion is a $\to$-formula, so we can conclude using $\to I$ again. This is as far as we get following this line of reasoning; we do the corresponding work for (7.3.19). It is a $\to$-formula so it should end with $\to I$. This is as far as we get in this case. We now know that the derivation can end as follows:

$$\dfrac{\dfrac{\dfrac{\bot}{\neg(\neg\varphi \wedge \neg\psi)}{\scriptstyle \to I}}{(\varphi \vee \psi) \to \neg(\neg\varphi \wedge \neg\psi)}{\scriptstyle \to I} \qquad \dfrac{\varphi \vee \psi}{\neg(\neg\varphi \wedge \neg\psi) \to (\varphi \vee \psi)}{\scriptstyle \to I}}{(\varphi \vee \psi) \leftrightarrow \neg(\neg\varphi \wedge \neg\psi)}{\scriptstyle \wedge I} \tag{7.3.20}$$

What is left is to derive $\bot$ from $\{\neg\varphi \wedge \neg\psi,\ \varphi \vee \psi\}$ as well as to derive $\varphi \vee \psi$ from $\neg(\neg\varphi \wedge \neg\psi)$. Let us start with the first one. The rule RAA is, according to Theorem 7.2.4 not needed at all here, so we look for an RAA-free derivation.

It is impossible to end with an introduction rule, since no such rule has the conclusion $\bot$. It thus has to end in an elimination rule. It is sufficient to look for a normal derivation. But then we know that the main premise in the last rule has to be a subformula of either $\neg\varphi \wedge \neg\psi$ or $\varphi \vee \psi$. Therefore, $\to E$, $\vee E$ and $\bot E$ are the only possibilities, but the last one is completely unnecessary to use in this case. Both $\to E$ and $\vee E$ are however possible to proceed with. We study the latter possibility. Then we get the following situation:

How does one know that a derivation of $\varphi \vee \psi$ from $\neg(\neg\varphi \wedge \neg\psi)$ cannot possible end with $\vee I$? Remember that φ and ψ stand for arbitrary formulas. If, for example, $\varphi = \psi = \top$ then one *can* in fact end with $\vee I$. But when we seek a derivation of $\varphi \vee \psi$ from $\neg(\neg\varphi \wedge \neg\psi)$ we look for a shape in which φ and ψ occur as symbols which can be replaced by arbitrary formulas. If the derivation ends with $\vee I$ then this means that we have to create a sub-derivation of φ from $\neg(\neg\varphi \wedge \neg\psi)$ (or of ψ), which should be correct no matter which formula we replace φ and ψ by. This is not possible. If we, for instance, put $\bot$ instead of φ and $\top$ instead of ψ, then a derivation from $\neg(\neg\varphi \wedge \neg\psi)$ to φ cannot be correct (according to the soundness theorem or the subformula property).

$$\dfrac{\dfrac{\dfrac{\dfrac{\varphi \vee \psi \qquad \bot \qquad \bot}{\bot}{\scriptstyle \vee E}}{\neg(\neg\varphi \wedge \neg\psi)}{\scriptstyle \to I}}{(\varphi \vee \psi) \to \neg(\neg\varphi \wedge \neg\psi)}{\scriptstyle \to I} \qquad \dfrac{\varphi \vee \psi}{\neg(\neg\varphi \wedge \neg\psi) \to (\varphi \vee \psi)}{\scriptstyle \to I}}{(\varphi \vee \psi) \leftrightarrow \neg(\neg\varphi \wedge \neg\psi)}{\scriptstyle \wedge I} \tag{7.3.21}$$

On the left, what remains to do is to derive $\bot$ from $\{\neg\varphi \wedge \neg\psi,\ \varphi\}$ respectively $\{\neg\varphi \wedge \neg\psi,\ \psi\}$. Then we have to use $\to E$, and at the top $\wedge E$. The left side is now done, and we proceed to complete the right side.

It is unreasonable to end a derivation of $\varphi \vee \psi$ with an introduction rule, since certain choices of φ and ψ makes them underivable. If we choose an elimination rule and seek a normal derivation, the main premise has to be a subformula of $\neg\varphi \wedge \neg\psi$, so $\vee E$ is excluded. The rules $\wedge E$ and $\to E$ are excluded since their main premises always contain the conclusion as a subformula. We are left with only $\bot E$ and RAA. The former is a dead end, according to what we have previously seen. Hence the only rule we are left with now is RAA. We have then the following situation:

$$\dfrac{\dfrac{\dfrac{\dfrac{[\varphi \vee \psi] \qquad \dfrac{\dfrac{[\neg\varphi \wedge \neg\psi]}{\neg\varphi}{\scriptstyle \wedge E} \quad [\varphi]}{\bot}{\scriptstyle \to E} \qquad \dfrac{\dfrac{[\neg\varphi \wedge \neg\psi]}{\neg\psi}{\scriptstyle \wedge E} \quad [\psi]}{\bot}{\scriptstyle \to E}}{\bot}{\scriptstyle \vee E}}{\neg(\neg\varphi \wedge \neg\psi)}{\scriptstyle \to I}}{(\varphi \vee \psi) \to \neg(\neg\varphi \wedge \neg\psi)}{\scriptstyle \to I} \qquad \dfrac{\dfrac{\bot}{\varphi \vee \psi}{\scriptstyle \mathrm{RAA}}}{\neg(\neg\varphi \wedge \neg\psi) \to (\varphi \vee \psi)}{\scriptstyle \to I}}{(\varphi \vee \psi) \leftrightarrow \neg(\neg\varphi \wedge \neg\psi)}{\scriptstyle \wedge I} \tag{7.3.22}$$

It remains to derive $\bot$ from $\{\neg(\neg\varphi \wedge \neg\psi),\ \neg(\varphi \vee \psi)\}$. We start solving this problem, once again, with the observation that we can look for RAA-free derivations.

In the next step we cannot use an introduction rule, and $\bot E$ is not applicable. Therefore, it is one the rules $\wedge E$, $\to E$ and $\vee E$ the ones we need to use. But in order for the derivation to be normal, the main premise must be a subformula of $\neg(\neg\varphi \wedge \neg\psi)$ or $\neg(\varphi \vee \psi)$, so the only possible rule we can use is $\to E$. The first main premise has to be $\neg\varphi$, $\neg\psi$, $\neg(\varphi \vee \psi)$ or $\neg(\neg\varphi \wedge \neg\psi)$. The first two cases are excluded, as they require that we derive the side premises φ respectively ψ, which is impossible, in general. The two other cases are both possible ways to proceed. We stop the process here, as it continues in a similar way. □

7.3.23 Exercise. Look for a derivation of $((P_1 \to P_1) \to P_1) \to P_1$.

7.3.24 Exercise (from the exam on 2006-01-12)**.** Look for a derivation of $((\varphi \to \psi) \to \psi) \leftrightarrow (\neg\varphi \to \psi)$.

7.3.25 Exercise. Look for a normal derivation from $P_1 \to P_2,\ \neg P_1 \to P_2$ to P_2.

7.3.26 Exercise (from the exam on 2005-10-20)**.** Derive $\neg(\neg\varphi \wedge (\psi \to \varphi)) \leftrightarrow (\varphi \vee \psi)$.

7.3.27 Exercise (from the exam on 2006-08-22)**.** Derive $(\neg P_1 \vee \neg P_2) \leftrightarrow \neg(P_1 \wedge P_2)$.

7.3.28 Exercise (from the exam on 2006-10-19)**.** Derive $\varphi \to ((\varphi \to \psi) \leftrightarrow \psi)$.

7.3.29 Exercise (from the exam on 2007-01-10)**.** Derive $((\varphi \vee \psi) \wedge \neg\varphi) \leftrightarrow \neg(\psi \to \varphi)$.

7.3.30 Exercise (from the exam on 2007-08-17)**.** Derive $((\varphi \to \psi) \wedge (\neg\varphi \to \psi)) \leftrightarrow \psi$.

7.3.31 Exercise (from the exam on 2007-10-18)**.** Derive $((\varphi \to \psi) \wedge \neg\psi) \leftrightarrow \neg(\varphi \vee \psi)$.

7.4 Summary

This chapter has dealt with normal derivations. The definition we have used is somewhat simplified compared to the usual one. You have seen that it leads to the *subformula property* for RAA-free normal derivations and other properties which simplify seeking normal derivations. The most important thing to bring with you for the rest of the course is the ability to efficiently find a derivation by searching for a normal derivation. It is also good to be able to prove that certain attempts to find derivations are doomed to fail. It is not important that you learn the theory of this section, but that you look at this section only as a way of helping you to look for a derivation.

Chapter 8

Completeness

We are now going to do something which seems completely impossible. We shall prove that everything which is true in every interpretation can be proved in natural deduction (if it can be expressed in the language we have built). The idea is the following: we shall construct an interpretation of the formulas in which the meaning of a formula will be precisely that it can be proven. Since they are true in this interpretation, they can be proven. This idea cannot be done in a straightforward way, but will be there as a guiding star. Instead of interpreting formulas so that they say of themselves that they can be proven, we interpret them in a way that their meaning is to be included in a so called *maximally consistent extension* of the set of undischarged assumptions. This eventually leads to the desired result. But first we need to define and study maximal consistency.

8.1 Maximal consistency

▶ **8.1.1 Definition.** Γ is *maximally consistent* if it is maximal amongst the consistent subsets of Form, ordered by inclusion. In simple words, this means that

1. Γ is consistent,
2. if $\Gamma \subseteq U \subseteq$ Form and U is consistent, then $U = \Gamma$.

Maximally consistent sets play more or less the same rules as maximal ideals do in ring theory.

An example of a maximally consistent set that one gets by considering a particular interpretation is the following: take every formula which is true in the interpretation.

None of the set of formulas we have considered so far are maximally consistent. In fact, every maximally consistent set is infinite, which follows from the fact that they are *closed under derivations* according to the next theorem.

8.1.2 Theorem (closure under derivations)**.** *If* Γ *is maximally consistent, and* $\Gamma \vdash \varphi$*, then* $\varphi \in \Gamma$.

Proof. Assume that Γ is maximally consistent and $\Gamma \vdash \varphi$. Let $U = \Gamma \cup \{\varphi\}$. Then $\Gamma \subseteq U$. It follows from the definition of maximal consistency that $\Gamma = U$ if we show that U is consistent. Assume therefore that $U \vdash \bot$. We will show that this leads to a contradiciton. Indeed, in that case, there should be a derivation of $\bot$ from the assumptions which are either φ or formulas in Γ. Since $\Gamma \vdash \varphi$, any assumptions of φ in the derivation can be replaced with derivations of φ from Γ. In this way we were able to construct a derivation of $\bot$ from Γ, which is impossible since Γ is consistent. □

8.1.3 Exercise. Show that every maximally consistent set of formulas is infinite.

8.1.4 Theorem. Γ *is maximally consistent if and only if it is consistent and whenever* $\Gamma \cup \{\varphi\}$ *is consistent, then* $\varphi \in \Gamma$.

Proof. (⇒) Take $U = \Gamma \cup \{\varphi\}$. If U is consistent, then according to the definition of maximal consistency, $U = \Gamma$, and hence $\varphi \in \Gamma$.

(⇐) Assume that $\Gamma \subseteq U \subseteq \text{Form}$ and that U is consistent. We will show that it follows then that $U = \Gamma$. Take an arbitrary formula $\varphi \in U$. Then $\Gamma \cup \{\varphi\} \subseteq U$, is consistent. Hence we have $\varphi \in \Gamma$. But φ was arbitrary in U, so we have $U \subseteq \Gamma$. Thus $U = \Gamma$. □

Since we have not seen a single example of a maximally consistent set, one can question why they are so important. The answer is that every consistent set Γ can be extended to a maximally consistent set Γ^* and it can be used to prove the completeness theorem, which is very useful. We shall start by constructing the extension and will prove later that it has the properties we want.

It is not important that you learn by heart how Γ^* is constructed, but you should understand the construction to be able to follow the reasoning in this chapter.

As a step along the way we start by constructing an infinite sequence of growing consistent subsets. A sequence is nothing else but a function from $\mathbb{N}$, so the definition is recursive and stated in two lines, as usual. We use the fact that Form is countable; that is to say, $\text{Form} = \{\varphi_0, \varphi_1, \varphi_2, \ldots\}$ for some enumeration of the formulas. We shall not explicitly define such an enumeration, but will content ourselves by asserting that it is possible to define it.

$$\Gamma_0 \stackrel{\text{def}}{=} \Gamma \tag{8.1.5}$$

$$\Gamma_{s(n)} \stackrel{\text{def}}{=} \begin{cases} \Gamma_n \cup \{\varphi_n\} & \text{if consistent} \\ \Gamma_n & \text{otherwise.} \end{cases} \tag{8.1.6}$$

Those who have concerns about definitions by cases which are difficult to decide may instead define $\Gamma_{s(n)}$ as $\Gamma_n \cup \{\varphi \mid \varphi = \varphi_n \text{ and } \Gamma_n \cup \{\varphi_n\} \text{ consistent}\}$.

8.1.7 Lemma. *The sequence $\{\Gamma_i\}$ is an increasing sequence of consistent sets if Γ is consistent.*

Proof. That the sequence is increasing follows from the fact that (8.1.6) specifies that the formulas which are in Γ_n will also be in $\Gamma_{s(n)}$. But we require more to show that every set in the sequence is consistent.

We will do the proof by induction. That Γ_0 is consistent follows from (8.1.5) and that Γ is consistent. Let us now do the induction step and assume, hence, the inductive hypothesis: that Γ_n is consistent. It follows immediately by construction of $\Gamma_{s(n)}$ that it is consistent. □

We now let

$$\Gamma^* = \bigcup_{i=0}^{\infty} \Gamma_i \,. \tag{8.1.8}$$

8.1.9 Lemma. *If $\Gamma^* \vdash \varphi$ then $\Gamma_n \vdash \varphi$ holds for some n.*

Proof. Assume that $\Gamma^* \vdash \varphi$; that is, there are $\gamma_1, \ldots, \gamma_m \in \Gamma^*$ and a derivation from $\gamma_1, \ldots, \gamma_m$ to φ. Since every γ_j belongs to Γ^*, which is the union of all such Γ_i, then every γ_j is in some Γ_i. Take n sufficiently large as to have every γ_j in Γ_n. □

8.1.10 Theorem. *If Γ is consistent, Γ^* is maximally consistent.*

Proof. Assume first that Γ^* is inconsistent; that is, that $\Gamma^* \vdash \bot$. Then, according to the previous lemma, $\Gamma_n \vdash \bot$ for some n, which is not the case, according to Lemma 8.1.7. Thus, Γ^* is consistent.

We shall now prove that Γ^* is maximally consistent through an application of Theorem 8.1.4. Assume therefore that $\Gamma^* \cup \{\varphi\}$ is consistent. Take n so that $\varphi = \varphi_n$. Then $\varphi \in \Gamma_{s(n)}$, according to (8.1.6). Hence, $\varphi \in \Gamma^*$. □

8.1.11 Theorem. *If $\varphi \notin \Gamma^*$, then $\neg\varphi \in \Gamma^*$.*

Proof. Assume that $\varphi \notin \Gamma^*$. We can show that $\neg\varphi \in \Gamma^*$ by applying Theorem 8.1.4 if we prove that $\Gamma^* \cup \{\neg\varphi\}$ is consistent. Assume therefore that it is inconsistent, and derive a contradiction.

We should then have some derivation

$$\begin{array}{c} \gamma_1 \ \cdots \ \gamma_n \quad \neg\varphi \\ \vdots \\ \bot \end{array} \tag{8.1.12}$$

where $\gamma_1, \ldots, \gamma_n \in \Gamma^*$. But we can then continue with RAA and discharge the assumption $\neg\varphi$. Then we would have $\Gamma^* \vdash \varphi$, and since maximally consistent sets are closed under derivations, we would have $\varphi \in \Gamma^*$, which contradicts our assumption. □

8.1.13 Exercise. Show that if $\neg\varphi \notin \Gamma^*$ then $\varphi \in \Gamma^*$.

8.1.14 Exercise. Show that if $\neg\psi \in \Gamma^*$ and $(\varphi \vee \psi) \in \Gamma^*$, then $\varphi \in \Gamma^*$.

8.1.15 Exercise. Show that if $\psi \in \Gamma^*$ then $(\varphi \to \psi) \in \Gamma^*$.

8.1.16 Exercise. Show that if $\varphi \notin \Gamma^*$, then $(\varphi \to \psi) \in \Gamma^*$.

8.1.17 Exercise (from the exam on 2004-01-08)**.**

a) Prove that $\{P_1, P_2, P_3, \neg P_1 \vee \neg P_2\}$ is inconsistent.

b) Is the set of all propositional variables maximally consistent?

8.2 Completeness

We will now use maximally consistent extensions to find interpretations in which all formulas in a consistent set are true.

8.2.1 Exercise. Show that if Γ has a model, then Γ is consistent.

> Remind yourself of the definition of a (4.2.39).

8.2.2 Lemma (model existence lemma)**.** *If Γ is consistent, then it has a model.*

Proof. Assume that Γ is consistent and that Γ^* is a maximally consistent extension of Γ. We define an interpretation by interpreting every propositional variable P_i as the proposition $P_i \in \Gamma^*$. We will check that for every $\varphi \in \text{Form}$ it holds that $\varphi \in \Gamma^* \iff [\![\varphi]\!] = 1$. In this case we know that every formula in Γ will have truth value 1, so the interpretation will be a model of Γ.

We prove this claim by induction on the complexity of the formula. There are a number a cases to consider– one for every sort of formulas.

For propositional variables, this is true by definition.

For $\top$, both $\top \in \Gamma^*$ and $[\![\top]\!] = 1$ hold.

For $\bot$ it holds that $\bot \notin \Gamma^*$ (since Γ^* is consistent) and $[\![\bot]\!] = 0$.

For formulas of the form $\varphi \wedge \psi$ we use the inductive hypothesis, which says that $\varphi \in \Gamma^* \iff [\![\varphi]\!] = 1$ and the same for ψ. If, therefore, $(\varphi \wedge \psi) \in \Gamma^*$, then it follows, since maximally consistent sets are closed under derivations, that $\varphi, \psi \in \Gamma^*$, and thus $[\![\varphi \wedge \psi]\!] = [\![\varphi]\!] \wedge [\![\psi]\!] = 1 \wedge 1 = 1$. Conversely, if $[\![\varphi \wedge \psi]\!] = 1$, then we have $[\![\varphi]\!] = [\![\psi]\!] = 1$, and hence $\varphi, \psi \in \Gamma^*$. Thus, $(\varphi \wedge \psi) \in \Gamma^*$ because Γ^* is closed under derivations.

For formulas of the form $\varphi \vee \psi$ we also use the induction hypothesis, but consider two cases: $[\![\psi]\!] = 0$ respectively $[\![\psi]\!] = 1$. In the first case we have, by inductive hypothesis, that $\psi \notin \Gamma^*$, which means, according to Theorem 8.1.11, that $\neg\psi \in \Gamma^*$. Assume now that $(\varphi \vee \psi) \in \Gamma^*$. Then we have, according to Exercise 8.1.14, that $\varphi \in \Gamma^*$, and hence $[\![\varphi]\!] = 1$ according to the inductive hypothesis, so it follows that $[\![\varphi \vee \psi]\!] = [\![\varphi]\!] \vee [\![\psi]\!] = 1 \vee 0 = 1$. On the other hand, if $[\![\varphi \vee \psi]\!] = 1$, it follows that $[\![\varphi]\!] = 1$ and thus by inductive hypothesis $\varphi \in \Gamma^*$ – and since Γ^* is closed under derivations, it follows that $(\varphi \vee \psi) \in \Gamma^*$.

The case where $\llbracket\psi\rrbracket = 1$ is easy: then $\psi \in \Gamma^*$ by inductive hypothesis. Hence we have both $(\varphi \lor \psi) \in \Gamma^*$ and $\llbracket\varphi \lor \psi\rrbracket \geq \llbracket\psi\rrbracket = 1$.

For formulas of the form $\varphi{\to}\psi$ we also use the inductive hypothesis. Assume that $(\varphi{\to}\psi) \in \Gamma^*$. We shall show that $\llbracket\varphi{\to}\psi\rrbracket = 1$, which means that if $\llbracket\varphi\rrbracket = 1$, then $\llbracket\psi\rrbracket = 1$. But if $\llbracket\varphi\rrbracket = 1$, the inductive hypothesis gives us $\varphi \in \Gamma^*$, and since Γ^* is closed under derivations, it follows that $\llbracket\psi\rrbracket = 1$, whereby, it follows from the inductive hypothesis that $\llbracket\psi\rrbracket = 1$. On the other hand, assume that $\llbracket\varphi \to \psi\rrbracket = 1$, we will show that $(\varphi \to \psi) \in \Gamma^*$. In the case $\llbracket\varphi\rrbracket = 1$, we must have $\llbracket\psi\rrbracket = 1$, and by inductive hypothesis it follows that $\psi \in \Gamma^*$, whereby Exercise 8.1.15 gives $(\varphi \to \psi) \in \Gamma^*$. In the case $\llbracket\varphi\rrbracket = 0$ we have $\varphi \notin \Gamma^*$ and hence by Exercise 8.1.16 we have $(\varphi \to \psi) \in \Gamma^*$ □

We are now ready to prove the converse of the soundness theorem.

8.2.3 Theorem (completeness)**.** $\Gamma \vDash \varphi \;\Rightarrow\; \Gamma \vdash \varphi$

Proof. Assume that $\Gamma \vDash \varphi$ and that $\Gamma \cup \{\neg\varphi\}$ is consistent. Then by the model existence lemma we have that $\Gamma \cup \{\neg\varphi\}$ has a model. But it is a model of Γ, and hence of φ, which contradicts that it is a model of $\neg\varphi$. Therefore, $\Gamma \cup \{\neg\varphi\}$ is inconsistent, so there must be a derivation of $\bot$ from $\gamma_1, \ldots, \gamma_n, \neg\varphi$. Then there is a derivation of φ from Γ:

$$\dfrac{\begin{array}{c}\gamma_1 \;\cdots\; \gamma_n \quad [\neg\varphi] \\ \vdots \\ \bot\end{array}}{\varphi}\,\text{RAA}$$

□

8.2.4 Exercise. Formulate the completeness theorem 8.2.3 in words. Compare to how Theorem 6.1.19 could be formulated in words in Theorem 6.1.5.

8.2.5 Exercise.

a) Show that a formula is *derivable* if and only if it is true in *all* interpretations; that is to say, if and only if *all* interpretations are models of the formula.

These are important and very useful principles.

b) Show that a set is *consistent* if and only if its formulas are true in *some common* interpretation; that is to say, if and only if *some* interpretation is a model of the formulas in the set.

8.2.6 Exercise. Show that $\varphi \leftrightarrow \psi$ can be derived if and only if $\varphi \approx \psi$.

8.2.7 Exercise (cf. exercises 5.6.6 and 6.1.29)**.** Show that $(\varphi \lor \psi) \to (\varphi \land \psi)$ is derivable if and only if $\varphi \approx \psi$.

8.2.8 Exercise. Show that if one substitutes a, b, c with P_1, P_2, P_3 in the Boolean axioms (Figure 1.1) and $=$ with $\leftrightarrow$, then the axioms become formulas which are derivable in natural deduction.

8.3 Summary

We have proved the converse of the soundness theorem. While the soundness theorem says that everything that is derivable is true in every interpretation, the completeness theorem says that everything which is true in every interpretation can be derived. It is thereby clear that the rules we have introduced in natural deduction are sufficient for our purposes: if something is not derivable from these rules, we would not like to derive it, since it is false in some interpretation. The most important thing to take with you for the rest of the course is the understanding of what completeness means, and how it can be used to show that a formula can be derived without actually constructing the derivation.

Part III

Predicate logic

Chapter 9

The language of predicate logic

The logic we have studied so far is called *propositional logic*, since deals with whole propositions and combine them to construct composite propositions. But this will not get us very far if we want to do mathematics. The problem is that one cannot express propositions such as “2 is even” in the language of propositional logic. The best one can do is symbolize it using a propositional variable. One then has to have another propositional variable symbolizing “3 is even” and a third one for “4 is even”, and so on. However, it would be better to have symbols in the language for 2, 3 and 4 and symbolize directly the *predicate* “is even”. One should also need to handle mathematical objects such as 2, 3, 4 and propositions about such objects. We shall do this by studying *predicate logic*.

Presumably you recognize the term *predicate* from grammar, which borrowed it from logic.

9.1 Terms

To refer to mathematical objects one uses *terms*. These are built from variables and function symbols. If, for instance, one would like to have terms to deal with numbers, one would need variables $x_0, x_1, x_2, \dots$ and function symbols for $+$ and $\cdot$. One also need symbols for 0 and 1, but instead of introducing a separate category of symbols for these, we consider them to be *nullary* function symbols, that is, functions that do not take any argument (they are *constants*). Such function symbols are therefore sometimes called *constant symbols*. As you can see, we can have different *arities* for our function symbols. We use $f_1, \dots, f_m$ as function symbols and denote their arities by $a_1, \dots, a_m$. If we would like f_1, f_2, f_3, f_4, f_5 to be interpreted as $0, 1, +, \cdot, -$, where $-$ is a unary *negation*, they should have the following arities:

a stands for arity.

$$a_1 = 0 \tag{9.1.1}$$
$$a_2 = 0 \tag{9.1.2}$$
$$a_3 = 2 \tag{9.1.3}$$
$$a_4 = 2 \tag{9.1.4}$$
$$a_5 = 1\,. \tag{9.1.5}$$

▶ **9.1.6 Definition.** Given an arity a_i for every function symbol f_i we define a set Term inductively by the following rules.

Compare to Chapter 3 and Definition 4.1.2.

$$\frac{i \in \mathbb{N}}{x_i \in \text{Term}}$$

$$\frac{t_1 \in \text{Term} \quad \cdots \quad t_{a_i} \in \text{Term}}{f_i(t_1, \dots, t_{a_i}) \in \text{Term}}$$

where we have a rule of the second type for every function symbol.

Instead of $f_i()$ we write f_i (nullary function symbols).

In our case, with the arities we chose above, we get:

$$\frac{}{f_1 \in \text{Term}} \tag{9.1.7}$$

$$\frac{}{f_2 \in \text{Term}} \tag{9.1.8}$$

$$\frac{t_1 \in \text{Term} \qquad t_2 \in \text{Term}}{f_3(t_1, t_2) \in \text{Term}} \tag{9.1.9}$$

$$\frac{t_1 \in \text{Term} \qquad t_2 \in \text{Term}}{f_4(t_1, t_2) \in \text{Term}} \tag{9.1.10}$$

$$\frac{t_1 \in \text{Term}}{f_5(t_1) \in \text{Term}} \tag{9.1.11}$$

In practice, it is not necessary to think too much about the arity type, since it is given by the context. On the other hand, it is important that one remembers that Term is not uniquely determined, but that it depends on the choice of arity type

The definition of Term depends on m and $a_1, \ldots, a_m$. It is, therefore, necessary to fix an *arity type* before one can start. This means precisely that one chooses $m, a_1, \ldots, a_m \in \mathbb{N}$. In the arity type, some other things should also be introduced, but we will wait a while before doing this. The final definition of arity type can be found in 9.2.1.

9.1.12 Example. Give the tree which proves that $f_4(f_5(f_2), x_1)$ is a term.

Solution.

$$\dfrac{\dfrac{\dfrac{}{f_2 \in \text{Term}}}{f_5(f_2) \in \text{Term}} \qquad \dfrac{\dfrac{\dfrac{}{0 \in \mathbb{N}}}{1 \in \mathbb{N}}}{x_1 \in \text{Term}}}{f_4(f_5(f_2), x_1) \in \text{Term}}$$

□

9.1.13 Exercise. Construct a tree which shows that $f_5(f_5(f_3(x_2, f_4(f_1, x_0))))$ is a term.

It turns out that it is important to use the notion of a variable *occurring* in a term. It means precisely what it seems, i.e., that when one reads the term, one finds the variable in it. The proper definition must be somewhat different, since we must define it according to the principle of inductively defined sets. It also has the advantage that limit cases become clearer. Would it, for instance, be correct to say that the variable x_1 occur in the term x_1? The answer is yes, as we choose it to be so. The definition splits, as usual, into cases given by the rules we have to construct terms.

▶ **9.1.14 Definition.** We define a variable as *occurring* in a term by

$$x_i \text{ occurs in } x_j \stackrel{\text{def}}{=} (i = j)$$
$$x_i \text{ occurs in } f_j(t_1, \ldots, t_{a_j}) \stackrel{\text{def}}{=} x_i \text{ occurs in some argument.}$$

With "some argument", we mean one of the terms $t_1, \ldots, t_{a_j}$. A special case is that of nullary function symbols: then x_i does not occur in any argument, and therefore no variable occurs in constant symbols. This fact is in acordance to what we usually look at things.

9.1.15 Exercise. Solve these exercises directly by looking at the variables in the terms and, more formally, by using the definition we just gave.

Note that, just by matching, the answer is *yes*. What is the correct answer?

a) Does x_2 occur in x_{23}?

b) Does x_0 occur in $f_4(x_0, x_1)$?

c) Does x_0 occur in $f_4(x_0, x_0)$?

d) Does x_2 occur in $f_3(x_0, f_1)$?

e) Does x_2 occur in $f_4(f_3(x_0, x_1), f_3(x_2, x_3))$?

We shall now define substitution. Substitution means that one "replaces a variable with some expression". For instance, we are used to "substituting" numbers such as 2 and 4 in an expression like "x^2", so that we get the expression "2^2", respectively "4^2". When substituting, it is always a *variable* that we substitute for. This is in accordance with the usual use of language. For instance, it is not usual "to substitute x for 2 in x^2". One can do substitutions in whole propositions, as when we are substituting to check the solution of an equation, but we shall start by substituting *terms* for *variables* in *terms*.

One can study substitution for other things than variables, in which case it is called *replacement* instead of substitution *substitution*. One can handle this by means of substitution, which we shall do in a while.

9.1.16 Definition (substitution of terms in terms)**.**

$$x_i[t/x_j] \stackrel{\text{def}}{=} \begin{cases} t & \text{if } j = i \\ x_i & \text{if } j \neq i \end{cases}$$
$$f_i(t_1, \ldots, t_{a_i})[t/x_j] \stackrel{\text{def}}{=} f_i(t_1[t/x_j], \ldots, t_{a_i}[t/x_j])\,.$$

Precedence rule: Substitution binds to the left, so $t[s/x_0][u/x_1]$ means $(t[s/x_0])[u/x_1]$.

9.1.17 Exercise. Compute

a) $f_3(x_0, f_1)[x_1/x_0]$

b) $f_3(x_0, x_1)[x_1/x_0][x_0/x_1]$

c) $f_3(x_0, f_1)[f_4(f_3(x_0, x_1), f_3(x_2, x_3))/x_2]$

9.1.18 Exercise. Show that if $s, t \in \text{Term}$ and x_j does not occur in t, then $t[s/x_j] = t$.

9.1.19 Exercise. Show that $t[x_i/x_j][x_j/x_i] = t$ does not always hold, but that it is true if x_i does not occur in t.

9.2 Formulas

We have already defined a set Form of formulas (Definition 4.1.2). We shall now modify this definition so that we can also have formulas that contain terms.

First of all, we generalize the idea of propositional variables. We now allow P_1, P_2, P_3 to take *arguments*, precisely as functions do. In the section about terms, we saw constant symbols as nullary function symbols, and the old propositional variables will now be seen as nullary *relation symbols*. We therefore need to have an arity for relation symbols. We denoted the arity for function symbols by $a_1, \ldots, a_m$, and we will now denote the arity for relation symbols by $r_1, \ldots, r_n$.

There are many names for relation symbols. Some prefer to call them predicate symbols.

9.2.1 Definition. By *arity type* we mean a list

$$\langle r_1, \ldots, r_n; a_1, \ldots, a_m \rangle\ ,$$

where $n, m, r_1, \ldots, r_n, a_1, \ldots, a_m \in \mathbb{N}$.

Remember that n or m, or both, could be 0.

Once an arity type is established, then also a *language* on that arity type is defined. It consists of *terms* which we introduced in the previous section, and of *formulas*, which we shall soon define. We need formulas to express *equality*; that is, propositions such as $1 \cdot x = x$. Since we want formulas which are easy to read, it is good to use a notation which ressembles $=$, but at the same time it is good to see the difference between formulas and interpretation, so we modify the notation a little and write $\doteq$. This means that $\doteq$ shall be seen as a *symbol* for $=$. We also need to express that something is true *for all* elements or *for some* element. We will do this with the symbols $\forall$, respectively $\exists$. In general, we imitate Definition 4.1.2.

When the symbol $\doteq$ is contained in an expression, it will remind that the expression is an element of Form. When the symbol $=$ is in an expression, it is an informal assertion, if anything. We can, for instance, write $\varphi = (x_0 \doteq x_0)$ to mean that φ is the formula $x_0 \doteq x_0$.

Instead of $P_i()$ we write P_i (nullary relation symbols).

While $\wedge$, $\vee$, $\to$, $\top$, $\bot$ are called *connectives*, one calls $\forall$ and $\exists$ *quantifiers*. The symbols were invented at a time when the easiest way to find new symbols was to use letters types already made, upside down.

9.2.2 Definition. Let $\langle r_1, \ldots, r_n; a_1, \ldots, a_m \rangle$ be an arity type. We define the set of formulas inductively as follows:

$$\frac{t_1 \in \text{Term} \qquad t_2 \in \text{Term}}{t_1 \doteq t_2 \in \text{Form}}$$

$$\frac{t_1 \in \text{Term} \ \cdots \ t_{r_i} \in \text{Term}}{P_i(t_1, \ldots, t_{r_i}) \in \text{Form}}$$

$$\frac{}{\top \in \text{Form}}$$

$$\frac{}{\bot \in \text{Form}}$$

$$\frac{\varphi \in \text{Form} \qquad \psi \in \text{Form}}{(\varphi \wedge \psi) \in \text{Form}}$$

$$\frac{\varphi \in \text{Form} \qquad \psi \in \text{Form}}{(\varphi \vee \psi) \in \text{Form}}$$

$$\frac{\varphi \in \text{Form} \qquad \psi \in \text{Form}}{(\varphi \to \psi) \in \text{Form}}$$

$$\frac{\varphi \in \text{Form} \qquad i \in \mathbb{N}}{\forall x_i \varphi \in \text{Form}}$$

$$\frac{\varphi \in \text{Form} \qquad i \in \mathbb{N}}{\exists x_i \varphi \in \text{Form}}$$

Note that the rule for $\doteq$ has precisely the same form as the rule for binary relation symbols. The only distinction will be that the interpretation of the latter can vary, while the interpretation of $\doteq$ will always be equality.

Remind yourself of Definition 4.1.3, page 30.

We keep Definition 4.1.3 for predicate logic as well: $\neg$ and $\leftrightarrow$ are regarded as defined operations.

9.2.3 Exercise. Let the arity type be $\langle 2, 3; 0, 1 \rangle$. Construct the tree which shows that $P_2(x_0, f_1, f_2(x_1)) \to \neg P_1(x_1, x_1)$ is a formula.

This exercise shows that propositional logic can be looked at as a special case of predicate logic, but that we have more formulas, even without the symbols for $\forall$ and $\exists$.

9.2.4 Exercise. Let the arity type be $\langle 0, 0, 0, 0; \rangle$.

a) Construct the tree which shows that $P_1 \wedge P_2 \to P_2 \wedge P_1$ is a formula.

b) Explain why every formula in predicate logic is a formula, even with our new definition of Form, if it does not contain any other predicate symbols than $P_1, \ldots, P_4$.

c) Give examples of two different terms (with their arity type).

d) Give examples of two different formulas (with their arity type) which were not formulas according to Definition 4.1.2. Construct the tree which shows that they are formulas according to Definition 9.2.2.

9.2.5 Definition. A formula is said to be *propositional* if it does not contain anything which was not part of propositional logic. Formally we define:

$$\begin{aligned}
(t_1 \doteq t_2) \text{ propositional} &\stackrel{\text{def}}{=} \text{false}\\
P_j(t_1,\ldots,t_{r_j}) \text{ propositional} &\stackrel{\text{def}}{=} (r_j = 0)\\
\top \text{ propositional} &\stackrel{\text{def}}{=} \text{true}\\
\bot \text{ propositional} &\stackrel{\text{def}}{=} \text{true}\\
(\varphi \wedge \psi) \text{ propositional} &\stackrel{\text{def}}{=} \varphi \text{ propositional and } \psi \text{ propositional}\\
(\varphi \vee \psi) \text{ propositional} &\stackrel{\text{def}}{=} \varphi \text{ propositional and } \psi \text{ propositional}\\
(\varphi \to \psi) \text{ propositional} &\stackrel{\text{def}}{=} \varphi \text{ propositional and } \psi \text{ propositional}\\
(\forall x_j \varphi) \text{ propositional} &\stackrel{\text{def}}{=} \text{false}\\
(\exists x_j \varphi) \text{ propositional} &\stackrel{\text{def}}{=} \text{false}
\end{aligned}$$

9.2.6 Exercise.

a) Show that $P_1 \wedge P_2 \to P_2 \wedge P_1$ is propositional according to the definition.

b) Show that your examples from Exercise 9.2.4 d are not propositional according to the definition.

9.2.7 Exercise. Define properly what "x_i occurs in φ" should mean, where $\varphi \in \text{Form}$.

> *Hint.* Think about the fact that all definitions shall be divided into cases according to the form of φ. Look at Definition 9.1.14 for inspiration.

9.2.8 Exercise. Show that if φ is a propositional formula, it is false that x_i occurs in φ.

We now reach the definition of substitution of terms for variables in formulas. Note particularly how the cases of $\forall$ and $\exists$ are handled. This may be somewhat surprising.

Precedence rule: $\forall$ and $\exists$ bind strongly, so $\forall x_0 \varphi \to \psi$ means $(\forall x_0 \varphi) \to \psi$. Substitution binds even stronger, so $\forall x_0 \varphi[t/x_0]$ means $\forall x_0(\varphi[t/x_0])$.

▶ **9.2.9 Definition** (substitution of terms in formulas)**.**

$$\begin{aligned}
(t_1 \doteq t_2)[t/x_j] &\stackrel{\text{def}}{=} (t_1[t/x_j] \doteq t_2[t/x_j])\\
P_i(t_1,\ldots,t_{r_i})[t/x_j] &\stackrel{\text{def}}{=} P_i(t_1[t/x_j],\ldots,t_{r_i}[t/x_j])\\
\top[t/x_j] &\stackrel{\text{def}}{=} \top\\
\bot[t/x_j] &\stackrel{\text{def}}{=} \bot\\
(\varphi_1 \wedge \varphi_2)[t/x_j] &\stackrel{\text{def}}{=} (\varphi_1[t/x_j] \wedge \varphi_2[t/x_j])\\
(\varphi_1 \vee \varphi_2)[t/x_j] &\stackrel{\text{def}}{=} (\varphi_1[t/x_j] \vee \varphi_2[t/x_j])\\
(\varphi_1 \to \varphi_2)[t/x_j] &\stackrel{\text{def}}{=} (\varphi_1[t/x_j] \to \varphi_2[t/x_j])\\
(\forall x_i \varphi)[t/x_j] &\stackrel{\text{def}}{=} \begin{cases} \forall x_i \varphi & \text{if } j = i\\ \forall x_i \varphi[t/x_j] & \text{if } j \neq i \end{cases}\\
(\exists x_i \varphi)[t/x_j] &\stackrel{\text{def}}{=} \begin{cases} \exists x_i \varphi & \text{if } j = i\\ \exists x_i \varphi[t/x_j] & \text{if } j \neq i \end{cases}
\end{aligned}$$

9.2.10 Exercise. Compute

a) $(x_1 \doteq x_2 \wedge P_1(f_1(x_1,x_2)))[f_2/x_1]$

b) $(x_1 \doteq x_2 \wedge \forall x_1(x_1 \doteq x_2))[f_2/x_1]$

c) $\forall x_1 \forall x_2(x_1 \doteq x_2 \wedge x_2 \doteq x_3)[x_3/x_2]$

If you have solved the exercises correctly (check the solutions) you will note that the variable one substitutes is *not* always replaced in every place it occurs. One replaces it only when its occurrence is "free". We shall sooon define what this means, but let us first consider an example.

As you probably remember, it is true that:

$$\int_0^1 x\,\mathrm{d}x = 1/2\,. \tag{9.2.11}$$

This also means that

$$x + \int_0^1 x\,\mathrm{d}x = x + 1/2\,. \tag{9.2.12}$$

This is a general formula, where we can substitute x for whatever we like. If we substitute x by 3, we conclude that:

$$3 + \int_0^1 x\,\mathrm{d}x = 3 + 1/2\,. \tag{9.2.13}$$

Note that we do *not* replace the x which is inside the integral. One says that the integral *binds* this x. While the x outside the integral is used as a symbol for an arbitrary number, the x inside the integral is used as an integration variable. In the same way, $\forall$ and $\exists$ bind variables. The variables which are not bound are called *free*. The exact definition is as follows:

▶ **9.2.14 Definition** (free variables)**.** We define "occurs freely in", or, more succinctly "free in" in the following way, where Definition 9.1.14 is used in some cases.

$$\begin{aligned}
x_i \text{ free in } (t_1 \doteq t_2) &\stackrel{\text{def}}{=} x_i \text{ occurs in either } t_1 \text{ or } t_2\\
x_i \text{ free in } P_j(t_1,\dots,t_{r_j}) &\stackrel{\text{def}}{=} x_i \text{ occurs in some of } t_1,\dots,t_{r_j}\\
x_i \text{ free in } \top &\stackrel{\text{def}}{=} \text{false}\\
x_i \text{ free in } \bot &\stackrel{\text{def}}{=} \text{false}\\
x_i \text{ free in } (\varphi \wedge \psi) &\stackrel{\text{def}}{=} x_i \text{ free in } \varphi \text{ or } x_i \text{ free in } \psi\\
x_i \text{ free in } (\varphi \vee \psi) &\stackrel{\text{def}}{=} x_i \text{ free in } \varphi \text{ or } x_i \text{ free in } \psi\\
x_i \text{ free in } (\varphi \to \psi) &\stackrel{\text{def}}{=} x_i \text{ free in } \varphi \text{ or } x_i \text{ free in } \psi\\
x_i \text{ free in } (\forall x_j \varphi) &\stackrel{\text{def}}{=} i \neq j \text{ and } x_i \text{ free in } \varphi\\
x_i \text{ free in } (\exists x_j \varphi) &\stackrel{\text{def}}{=} i \neq j \text{ and } x_i \text{ free in } \varphi
\end{aligned}$$

Did you understand the difference between a variable occurring in φ and a variable occurring freely in φ?

9.2.15 Exercise. That a variable occurs bound in a formula means that it is bound by $\forall$ or $\exists$. Give a recursive definition in the same spirit as the previous one.

9.2.16 Exercise.

a) Does x_1 occur freely in $x_1 \doteq x_2$?

b) Does x_1 occur freely in $x_1 \doteq x_1$?

c) Does x_1 occur freely in $(x_1 \doteq x_2 \wedge P_1(f_1(x_1, x_2)))$?

d) Does x_1 occur freely in $\forall x_1(x_1 \doteq x_2))$?

e) Does x_2 occur freely in $\forall x_1(x_1 \doteq x_2))$?

f) Does x_1 occur freely in $(x_1 \doteq x_2 \wedge \forall x_1(x_1 \doteq x_2))$?

g) Does x_2 occur freely in $\forall x_1 \forall x_2(x_1 \doteq x_2 \wedge x_2 \doteq x_3)$?

h) Does x_1 occur freely in $\neg(x_1 \doteq x_1)$?

Sometimes we write $x_i \in \mathrm{FV}(\varphi)$ instead of "x_i occurs free in φ". In other words: $\mathrm{FV}(\varphi)$ is the set of free variables in φ.

9.2.17 Exercise.

a) Determine $\mathrm{FV}(x_1 \doteq x_2)$.

b) Determine $\mathrm{FV}(x_1 \doteq x_2 \wedge P_1(f_1(x_1, x_2)))$.

c) Determine $\mathrm{FV}(\forall x_1 \forall x_2(x_1 \doteq x_2 \wedge x_2 \doteq x_3))$.

d) Determine $\mathrm{FV}(\top)$.

e) Determine $\mathrm{FV}(\varphi \wedge \psi)$ if $\mathrm{FV}(\varphi) = \{x_1\}$ and $\mathrm{FV}(\psi) = \emptyset$.

f) Determine $\mathrm{FV}(\varphi \vee \psi)$ if $\mathrm{FV}(\varphi) = \{x_1\}$ and $\mathrm{FV}(\psi) = \emptyset$.

9.2.18 Exercise. Show that if x_j does not occur freely in φ then $\varphi[t/x_j] = \varphi$. Use an inductive proof (induction on the complexity of the formula).

9.2.19 Exercise (from the exam on 2003-01-09)**.** Consider the following formula:

$$\forall x_2(\forall x_1 P_1(x_1, x_2) \to \exists x_2(f_1(x_1) \doteq f_2(x_2, x_3))) \vee \forall x_3 \neg(x_1 \doteq x_3)\,.$$

Call this formula φ.

a) Determine $\mathrm{FV}(\varphi)$.

b) Perform the substitutions $\varphi[f_1(x_3)/x_1]$, $\varphi[x_1/x_2]$, $\varphi[f_2(x_1, x_3)/x_3]$.

9.2.20 Exercise (from the exam on 2005-01-07)**.** In this exercise, φ denotes formulas in the language of arity type $\langle 1; 1, 0\rangle$. That a formula is atomic means that it does not contain connectives ($\top$, $\bot$, $\wedge$, $\vee$, $\to$) or quantifiers ($\forall$, $\exists$).

a) Give examples of three different atomic formulas φ without any free variables.

b) Give examples of three different atomic formulas φ which satisfy $\mathrm{FV}(\varphi) = \{x_0, x_1\}$.

9.2.21 Exercise. Prove by induction that $t[x_i/x_i] = t$ and $\varphi[x_i/x_i] = \varphi$.

9.2.22 Exercise. Show that $\varphi[y/x][x/y] = \varphi$ if y does not occur in φ.

9.2.23 Exercise. Show that $\varphi[y/x][x/y] \neq \varphi$ can be true even when y does not occur freely in φ.

Hint. Take $\varphi = \forall x_0(x_1 \doteq x_1)$, $y = x_0$, $x = x_1$.

9.3 Summary

We have introduced the language of predicate logic in a way that follows closely the related development for propositional logic, but with the major difference that we now have two ingredients: terms and formulas. The formulas that do not contain any terms were now recognized as propositional formulas. Since we have introduced terms, we also needed to introduce a machinery to manipulate them: substitution. Predicate logic becomes much more complicated than propositional logic precisely because of substitution, but it has also many more applications. The most important thing to remember for the rest of the course is the knowledge of what exactly the sets Term and Form contain and how this depend on an *arity type*. It is also very important to know precisely how substitution is done and to know what it means for a variable to occur freely in a formula.

Chapter 10

Semantics

In this chapter we will, to a large extent, repeat what we have already done for propositional logic. However, we need to make some modifications to adjust to the more advanced situation we now have.

10.1 Interpretation of terms and formulas

To define an interpretation $\mathcal{A}$ it is not sufficient to choose propositions as interpretations for $P_1, P_2, \ldots$, since these symbols are no necessarilly nullary anymore, as they could now take arguments. Therefore, they shall be instead interpreted as relations. For instance, P_1 can be interpreted as $\leq$ if it takes two arguments. An interpretation $\mathcal{A}$ consists, more precisely, of the following:

- A *set* $|\mathcal{A}|$ which is called *domain* (of individuals); we think about it as the set of the elements about which the language speaks.
- For every relation symbol P_j, an r_j-ary relation $P_j^{\mathcal{A}}$ on $|\mathcal{A}|$. This means that $P_j^{\mathcal{A}}(b_1, \ldots, b_{r_j})$ is a *proposition*, which is true or false for every choice of $b_1, \ldots, b_{r_j} \in |\mathcal{A}|$.
- For every function symbol f_j, an a_j-ary function $f_j^{\mathcal{A}}$ on $|\mathcal{A}|$. This means that $f_j^{\mathcal{A}}(b_1, \ldots, b_{a_j})$ is an *element* in $|\mathcal{A}|$ for every choice of $b_1, \ldots, b_{a_j} \in |\mathcal{A}|$.
- A *valuation* of the variables, which is a function v from the variables to $|\mathcal{A}|$.

Notice that the set $|A|$ will always be non-empty, since we require the existence of a function from the non-empty set of variables into $|A|$.

10.1.1 Exercise. What special cases do we get when we interpret a nullary relation symbol or function symbol?

We will return to the usefulness of valuations below. For now you do not have to care very much about them

10.1.2 Example. Assume that we have a language of arity type $\langle ; 0, 1, 2, 2\rangle$ and we would like to interpret it involving natural numbers. We let

$$|\mathcal{A}| \stackrel{\text{def}}{=} \mathbb{N} \tag{10.1.3}$$
$$f_1^{\mathcal{A}} \stackrel{\text{def}}{=} 0 \tag{10.1.4}$$
$$f_2^{\mathcal{A}} \stackrel{\text{def}}{=} s \tag{10.1.5}$$
$$f_3^{\mathcal{A}} \stackrel{\text{def}}{=} + \tag{10.1.6}$$
$$f_4^{\mathcal{A}} \stackrel{\text{def}}{=} \cdot \tag{10.1.7}$$

and define also a valuation of the variables. Often we wait to decide the valuation until some concrete problem is solved. The reason for this is that most of the things we do do not depend on the valuation, so it is often not necessary to specify it.

The above definition of an interpretation is formulated more simply by saying that we interpret in the structure:

$$\langle \mathbb{N};;0,s,+,\cdot\rangle\,. \tag{10.1.8}$$

Notice that in the notation for structure, there is a semicolon more than in the notation for arity type. To the left of the first one writes the name of the domain. There is nothing corresponding to this in arity types. To the right of this one writes the relations one uses in the interpretation. Finally, in the last space one writes the functions.

A structure is therefore nothing more than a set together with relations and functions. The advantage of the notation (10.1.8) is that one can define the whole interpretation in one row. The ordering in (10.1.8) is relevant to be able to know what is the interpretation of symbols.

We assume in the sequel that we have a given arity type and a given interpretation $\mathcal{A}$ of the language. Let us look closer at what we shall use our *valuations* for.

Already in high school mathematics one states things as, for example, that $2^x = x^2$ does not hold *for all values of* x. Thus, one speaks of giving *values* to the variables. This is what the valuation does.

10.1.9 Example. An example of a valuation where the variables are the natural numbers is given by:

$$v(x_i) = i\,,$$

i.e. x_0 is given the value 0, while x_1 is given the value 1, and so on.

10.1.10 Example. A valuation can also be given by an infinite list. Let, for example, v be defined by

$$\begin{aligned} x_0 &\mapsto 7\\ x_1 &\mapsto 3\\ x_2 &\mapsto 7\\ x_3 &\mapsto 3\\ x_4 &\mapsto 7\\ &\vdots \end{aligned}$$

When we have a valuation of the variables we can recursively extend it to Term precisely as we did with Form in propositional logic. In this way, all terms are given a value in the domain.

10.1.11 Definition. Let

$$\begin{aligned} [\![x_i]\!] &\stackrel{\text{def}}{=} v(x_i)\\ [\![f_i(t_1,\ldots,t_{a_i})]\!] &\stackrel{\text{def}}{=} f_i^{\mathcal{A}}([\![t_1]\!],\ldots,[\![t_{a_i}]\!])\,. \end{aligned}$$

Note that this definition also depends on the interpretation $\mathcal{A}$, so it becomes necessary to write $[\![\varphi]\!]^{\mathcal{A}}$ when we need to specify that it is the interpretation $\mathcal{A}$ we have in mind.

10.1.12 Example. Interpret the language of arity type $\langle;0,1,2,2\rangle$ in the structure

$$\langle \mathbb{N};;0,s,+,\cdot\rangle \tag{10.1.13}$$

but leave the valuation of the variables unspecified. Compute the expression $[\![f_3(f_4(f_2(f_1),x_0),x_1)]\!]$ as far as possible.

Solution.

$$\begin{aligned} [\![f_3(f_4(f_2(f_1),x_0),x_1)]\!] &\stackrel{\text{def}}{=} f_3^{\mathcal{A}}(f_4^{\mathcal{A}}(f_2^{\mathcal{A}}(f_1^{\mathcal{A}}),[\![x_0]\!]),[\![x_1]\!])\\ &\stackrel{\text{def}}{=} s(0)\cdot v(x_0)+v(x_1)\\ &= v(x_0)+v(x_1). \end{aligned}$$

□

The answer is thus a function of the valuation of x_0 and x_1. If we let $x = v(x_0)$ and $y = v(x_1)$ we can answer that $[\![f_2(f_3(f_2(f_1), x_0), x_1)]\!] = x + y$. But we cannot calculate any further if we do not know the valuation of x_0 and x_1, that is to say, if we do not know more about the valuation of the variables.

We shall also give values to formulas, but to do this we need a technical detail. We will have to say things such as "the same interpretation as $\mathcal{A}$, but with x_3 given the value 7 instead". This interpretation we denote by $\mathcal{A}[x_3 \mapsto 7]$. The definition looks like this:

▶ **10.1.14 Definition** (reevaluation)**.** Let $\mathcal{A}$ be an interpretation whose valuation we denote by v. We then let

$$v[x_i \mapsto a](x_j) \stackrel{\text{def}}{=} \begin{cases} a & \text{if } i = j \\ v(x_j) & \text{otherwise} \end{cases}$$

and $\mathcal{A}[x_i \mapsto a]$ is the interpretation $\mathcal{A}$ but with the valuation v replaced by $v[x_i \mapsto a]$.

10.1.15 Example. If v is like in Example 10.1.10, then $v[x_1 \mapsto 0]$ is the same except in the case x_1:

$$\begin{aligned} x_0 &\mapsto 7 \\ x_1 &\mapsto 0 \\ x_2 &\mapsto 7 \\ x_3 &\mapsto 3 \\ x_4 &\mapsto 7 \\ &\vdots \end{aligned}$$

10.1.16 Example. If $a \in |\mathcal{A}|$, then

We do not have to know anything about v to compute this.

$$[\![x_0]\!]^{\mathcal{A}[x_0 \mapsto a]} = v[x_0 \mapsto a](x_0) = a.$$

10.1.17 Exercise. Simplify

a) $\mathcal{A}[x_i \mapsto a][x_i \mapsto b]$

b) $\mathcal{A}[x_i \mapsto [\![x_i]\!]^{\mathcal{A}}]$

c) $\mathcal{A}[x_i \mapsto [\![x_i]\!]^{\mathcal{A}[x_i \mapsto b]}]$

10.1.18 Exercise. Show that if $i \neq j$, then

$$\mathcal{A}[x_i \mapsto a][x_j \mapsto b] = \mathcal{A}[x_j \mapsto b][x_i \mapsto a]$$

, but that it is not necessarily the case if $i = j$. Show that in this case both sides of the equation can be simplified.

We can now define the truth values on Form:

▶ **10.1.19 Definition.** Let $\mathcal{A}$ be an interpretation. The truth values of formulas are given by:

$$
\begin{aligned}
[\![t_1 \doteq t_2]\!] &\stackrel{\text{def}}{=} \begin{cases} 1 & \text{if } [\![t_1]\!] = [\![t_2]\!] \text{ is true,} \\ 0 & \text{otherwise.} \end{cases} \\
[\![P_i(t_1, \ldots, t_{r_i})]\!] &\stackrel{\text{def}}{=} \begin{cases} 1 & \text{if } P_i^{\mathcal{A}}([\![t_1]\!], \ldots, [\![t_{r_i}]\!]) \text{ is true,} \\ 0 & \text{otherwise.} \end{cases} \\
[\![\top]\!] &\stackrel{\text{def}}{=} 1 \\
[\![\bot]\!] &\stackrel{\text{def}}{=} 0 \\
[\![\varphi \wedge \psi]\!] &\stackrel{\text{def}}{=} [\![\varphi]\!] \wedge [\![\psi]\!] \\
[\![\varphi \vee \psi]\!] &\stackrel{\text{def}}{=} [\![\varphi]\!] \vee [\![\psi]\!] \\
[\![\varphi \to \psi]\!] &\stackrel{\text{def}}{=} [\![\varphi]\!] \to [\![\psi]\!] \\
[\![\forall x_i \varphi]\!] &\stackrel{\text{def}}{=} \begin{cases} 1 & \text{if } [\![\varphi]\!]^{\mathcal{A}[x_i \mapsto a]} = 1 \text{ is true for } \textit{all } a \in |\mathcal{A}|, \\ 0 & \text{otherwise.} \end{cases} \\
[\![\exists x_i \varphi]\!] &\stackrel{\text{def}}{=} \begin{cases} 1 & \text{if } [\![\varphi]\!]^{\mathcal{A}[x_i \mapsto a]} = 1 \text{ is true for } \textit{some } a \in |\mathcal{A}|, \\ 0 & \text{otherwise.} \end{cases}
\end{aligned}
$$

When the interpretation $\mathcal{A}$ is implicit, we write $[x_i \mapsto a]$ instead of $\mathcal{A}[x_i \mapsto a]$, $[x_i \mapsto a][x_j \mapsto b]$ instead of $\mathcal{A}[x_i \mapsto a][x_j \mapsto b]$, etc.

10.1.20 Example. Compute $[\![\forall x_0(x_0 \doteq x_0)]\!]$.

Solution. According to the definition of $[\![\forall x_i \varphi]\!]$ we shall compute $[\![x_0 \doteq x_0]\!]^{[x_0 \mapsto a]}$ for all a, and investigate if the answer is always 1. According to the definition of $[\![x_0 \doteq x_0]\!]^{[x_0 \mapsto a]}$, this is 1 if $[\![x_0]\!]^{[x_0 \mapsto a]} = [\![x_0]\!]^{[x_0 \mapsto a]}$ is true. But it is, since $=$ is reflexive. Thus, the answer is 1. □

10.1.21 Exercise. Compute $[\![\forall x_0(x_0 \doteq x_1)]\!]$ and $[\![\exists x_0(x_0 \doteq x_1)]\!]$ as far as possible.

10.1.22 Exercise (from the exam on 2004-08-17)**.** Decide if $\exists x(P_1(x) \vee P_2(x)) \to (\exists x P_1(x) \vee \exists x P_2(x))$ is true in all interpretations.

If $[\![\varphi]\!]^{\mathcal{A}} = 1$ one says that φ is true in $\mathcal{A}$ and writes $\mathcal{A} \vDash \varphi$. If φ does not contain any free variables, $[\![\varphi]\!]^{\mathcal{A}}$ does not depend at all on the valuation, which explains that it is often unnecesasary to specify which valuation we are using. This fact follows from the following two theorems:

We use the symbols $x, y, z, \ldots$ as metavariables for object variables. This means that these symbols stand for *arbitrary* variables $x_0, x_1, x_2, \ldots$. We can never have $x_0 = x_1$, since these two symbols are different in Term, though we can have $x = y$, which means that x and y symbolize the same variable, for instance x_0.

10.1.23 Theorem. *If x does not occur in t, then $[\![t]\!]^{[x \mapsto a]} = [\![t]\!]$.*

Proof idea. Intuitively, as soon as one understands the symbolisms, this is an obvious consequence of what reevaluation means. What it does is to change the values of the variables. What the theorem says is just that if the variable whose value is changed does not occur in t, the value of t will not be changed. □

The formal proof shows how the machinery we have built works.

Proof. The proof is done by induction on the structure of terms, since we will show that something is true for *all* terms. The term t may either be of the form x_i or of the form $f_i(t_1, \ldots, t_{a_i})$. In the first case, we know, since x does not occur in t, that $x \neq x_i$. Therefore $[\![t]\!]^{[x \mapsto a]} = [\![x_i]\!]^{[x \mapsto a]} = [\![x_i]\!]$. If $t = f_i(t_1, \ldots, t_{a_i})$ holds, since x does not occur in t, we know that x does not occur in any of the arguments. The inductive hypothesis gives us $[\![t_j]\!]^{[x \mapsto a]} = [\![t_j]\!]$, so it follows that

$$
\begin{aligned}
[\![t]\!]^{[x \mapsto a]} &= [\![f_i(t_1, \ldots, t_{a_i})]\!]^{[x \mapsto a]} && (10.1.24) \\
&= f_i^{\mathcal{A}}([\![t_1]\!]^{[x \mapsto a]}, \ldots, [\![t_{a_i}]\!]^{[x \mapsto a]}) && (10.1.25) \\
&= f_i^{\mathcal{A}}([\![t_1]\!], \ldots, [\![t_{a_i}]\!]) && (10.1.26) \\
&= [\![f_i(t_1, \ldots, t_{a_i})]\!] = [\![t]\!]\,. && (10.1.27)
\end{aligned}
$$

□

10.1.28 Theorem. *If x does not occur freely in φ then* $[\![\varphi]\!]^{[x\mapsto a]} = [\![\varphi]\!]$.

Proof idea. Here we can also understand the theorem in an informal way. Since reevaluation is defined in such a way that it only changes the values of the *free* variables, it is clear that if x does not occur freely in φ, the value of φ will not change if we change the value of x. □

The task of the formal proof is, in this case, to exhibit the machinery, but also to actually check that everything works. It is a quite complex definition the one we have made, and it is not entirely obvious to see that it does exactly what we want to. The proof consists of checking that the definition works.

Proof. Also in this case we use an inductive proof, but there are many more cases since formulas can be constructed in many ways. We will also strengthen the theorem somewhat, to get a stronger inductive hypothesis: we will show that if x does not occur freely in φ, then $[\![\varphi]\!]^{\mathcal{A}[x\mapsto a]} = [\![\varphi]\!]^{\mathcal{A}}$ for *all* interpretations $\mathcal{A}$. This strengthening does not result in any differences in the use of the theorem, but rather changes its logical form. It has the advantage that we get as inductive hypothesis that the theorem holds for subformulas in *all* interpretations.

If φ is of the form $t_1 \doteq t_2$ or $P_i(t_1,\ldots,t_{r_i})$, the result follows from the previous theorem. In the other cases, except the quantifiers, the result follows quite directly after using the inductive hypothesis. We show here the case of $\forall$-formulas ($\exists$-formulas are handled similarly).

We shall show that if x does not occur freely in $\forall x_i\psi$, then

$$[\![\forall x_i\psi]\!]^{[x\mapsto a]} = [\![\forall x_i\psi]\!]\,. \tag{10.1.29}$$

That x does not occur freely in $\forall x_i\psi$ means that either $x = x_i$ or that x does not occur freely in ψ (see Definition 9.2.14). We split the proof into these two cases. It is sufficient to check that each of the sides in (10.1.29) are simultaneously 1 ; that is, that $[\![\psi]\!]^{[x\mapsto a][x_i\mapsto b]} = 1$ for all $b \in |\mathcal{A}|$ if and only if $[\![\psi]\!]^{[x_i\mapsto b]} = 1$ for all $b \in |\mathcal{A}|$.

If $x = x_i$ then $[\![\psi]\!]^{[x\mapsto a][x_i\mapsto b]}$ is simplified to $[\![\psi]\!]^{[x_i\mapsto b]}$ (Exercise 10.1.17 a), from which the result follows immmediately. If x does not occur freely in ψ we have, after using the inductive hypothesis on $\mathcal{A}[x_i \mapsto b]$, that $[\![\psi]\!]^{[x_i\mapsto b]} = [\![\psi]\!]^{[x_i\mapsto b][x\mapsto a]}$, and hence it is enough to prove that

$$[\![\psi]\!]^{[x\mapsto a][x_i\mapsto b]} = [\![\psi]\!]^{[x_i\mapsto b][x\mapsto a]}\,. \tag{10.1.30}$$

This is actually not true in general, but we have already handled the case where $x = x_i$ above, so we can now assume that $x \neq x_i$. Then, (10.1.30) follows from Exercise 10.1.18. □

10.1.31 Exercise (from the exam on 2005-01-07)**.** An *equivalence relation* is a binary relation $\sim$ which has the following properties for all a, b, c:

$$a \sim a \qquad \text{(reflexivity)}$$
$$\text{if } a \sim b, \text{ then } b \sim a \qquad \text{(symmetry)}$$
$$\text{if } a \sim b \text{ and } b \sim c, \text{ then } a \sim c\,. \qquad \text{(transitivity)}$$

Formalize these rules; that is, give three formulas $\gamma_r, \gamma_s, \gamma_t$ which represent these three rules. Choose a suitable arity type. The formulas should not contain free variables.

10.2 Models and countermodels

We shall introduce the notion of *model* also in predicate logic, as well as the notion of *countermodel*. A model is an interpretation in which one or more specified formulas are true, while a countermodel is an interpretation in which not all of the specified formulas are true.

10.2.1 Definition.

1. A model of φ is an interpretation $\mathcal{A}$ in which φ is true: $[\![\varphi]\!]^{\mathcal{A}} = 1$.

2. A model of a set Γ of formulas is a model of *all* formulas in Γ.

3. We say that Γ gives φ ($\Gamma \vDash \varphi$) if all models of Γ are models of φ.

4. We say that $\gamma_1, \ldots, \gamma_n \vDash \varphi$ if $\{\gamma_1, \ldots, \gamma_n\} \vDash \varphi$.

5. A countermodel of φ is an interpretation $\mathcal{A}$ in which φ is false, i.e., where $[\![\varphi]\!]^{\mathcal{A}} = 0$.

6. A countermodel for a set Γ of formulas is a countermodel for *at least one* formula in Γ.

Note that a countermodel for a set does not necessarily has to be a countermodel for *every* formula in the set. If an interpretation is *not* a model for the set, it is a countermodel.

Note the special case "$\vDash \varphi$". It can be conceived as having an empty list to the left of $\vDash$, so that it means the same as $\emptyset \vDash \varphi$, which is to say that all models of $\emptyset$ are models of φ. Since every interpretation is a model of $\emptyset$ (since it is a model of every formula in $\emptyset$), $\emptyset \vDash \varphi$ means that every interpretation is a model of φ. Thus, $\vDash \varphi$ is a way of expressing that φ is true in every interpretation. We then say that φ is a tautology, precisely as in propositional logic.

10.2.2 Exercise. Show that $\forall x\varphi \vDash \varphi$.

10.2.3 Exercise. Show that $\varphi \vDash \exists x\varphi$.

10.2.4 Example. Show that $\nvDash (x_0 \doteq x_1)$.

Proof. We will find a countermodel. The naive argument is: "let x_0 and x_1 have different values!". This works, but let us be more precise for the sake of practice. Choose, then, a structure, say the natural numbers, and let $v(x_i) = i$. We then get that

$$[\![x_0 \doteq x_1]\!] = 1 \iff [\![x_0]\!] = [\![x_1]\!] \iff 0 = 1$$

and since the last equation is actually false, we know that $[\![x_0 \doteq x_1]\!] = 0$. We have then a countermodel. □

10.2.5 Example. Show that if $t_1, t_2, t_3 \in \text{Term}$, then

1. $\vDash (t_1 \doteq t_1)$,
2. $(t_1 \doteq t_2) \vDash (t_2 \doteq t_1)$,
3. $(t_1 \doteq t_2), (t_2 \doteq t_3) \vDash (t_1 \doteq t_3)$.

In practice, it almost always works to write $[\![\varphi]\!]$ instead of $[\![\varphi]\!]^{\mathcal{A}}$ when computing, even if we consider more than one interpretation at the same time. Indeed, when we often work with an arbitrary interpretation with a certain property, one can just assume that the interpretation in question has that property and then compute as usual. Example 10.2.5 illustrates this.

Proof.
1. We will show that in all interpretations we have $[\![t_1 \doteq t_1]\!] = 1$. According to the definition, we need to show that we have $[\![t_1]\!] = [\![t_1]\!]$ in all interpretations, which we have, since $=$ is reflexive.

2. We shall show that if $[\![t_1 \doteq t_2]\!] = 1$ then we also have $[\![t_2 \doteq t_1]\!] = 1$. Assume therefore that $[\![t_1 \doteq t_2]\!] = 1$, which means that $[\![t_1]\!] = [\![t_2]\!]$ is true. Since $=$ is symmetric, it follows that $[\![t_2]\!] = [\![t_1]\!]$ is true. Hence, $[\![t_2 \doteq t_1]\!] = 1$.

3. Similarly to the previous item, but use now that $=$ is transitive.

□

10.2.6 Example. Show that there is $\varphi \in \text{Form}$ such that $\varphi \nvDash \forall x_0\varphi$.

Proof. Take for instance $\varphi = (x_0 \doteq x_1)$. Then $[\![\varphi]\!]^{\mathcal{A}} = 1$ if $[\![x_0]\!]^{\mathcal{A}} = [\![x_1]\!]^{\mathcal{A}}$. Let us therefore consider an interpretation $\mathcal{A}$ where this is the case. We have $[\![\forall x_0\varphi]\!] = 1$ only if $[\![\varphi]\!]^{\mathcal{A}[x_0 \mapsto a]} = 1$ for all a. But $[\![\varphi]\!]^{\mathcal{A}[x_0 \mapsto a]} = 1$ precisely when $a = [\![x_1]\!]^{\mathcal{A}}$, which is not true for all a in structures with more than one element.

To sum up: if we consider a structure with at least two elements and give x_0 and x_1 the same value, we have $[\![\varphi]\!] = 1$, but $[\![\forall x_0\varphi]\!] = 0$. □

10.2.7 Exercise. Show that it is not necessarily true that $\forall x_0(\varphi \vee \psi) \vDash \forall x_0\varphi \vee \psi$. In other words, that it is false for some choice of $\varphi, \psi \in \text{Form}$.

Hint. Let $\psi = \neg\varphi$ and use the same idea as in the previous example.

10.2.8 Exercise. Show that it is not necessarily true that $\forall x_0 \varphi \vee \psi \vDash \forall x_0(\varphi \vee \psi)$.

Hint. Let $\varphi = \bot$. Then use Example 10.2.6.

10.2.9 Exercise (from the exam on 2005-01-07)**.** Interpret the formula $\forall x_0 \exists x_1 (f_1(x_1, x_1) \doteq x_0)$ in the following structures and find its truth value in each one of them:

a) $\langle \mathbb{R};;+,0\rangle$

b) $\langle \mathbb{R};;\cdot,1\rangle$

c) $\langle \mathbb{C};;\cdot,1\rangle$

Here $\mathbb{R}$ are the real numbers and $\mathbb{C}$ are the complex numbers.

10.2.10 Exercise (from the exam on 2005-08-23)**.** Interpret the formula $\forall x_0 \exists x_1 P_1(x_0, x_1)$ in the following structures and find its truth value in each one of them:

a) $\langle (0,1);<;\rangle$

b) $\langle [0,1];<;\rangle$

Here $(0,1)$ is the open (real) interval between 0 and 1, while $[0,1]$ is the closed interval.

10.2.11 Exercise (from the exam on 2004-08-17)**.** Interpret the formula $\forall x_0 \forall x_1 (P_1(x_0, x_1) \rightarrow P_1(f_1(x_0, f_2), f_1(x_1, f_2)))$ in the following structures and find its truth value in each one of them:

a) $\langle \mathbb{R};\leq;+,1\rangle$

b) $\langle \mathbb{R};\leq;\cdot,-1\rangle$

c) $\langle \mathbb{R};\neq;\cdot,0\rangle$

Here $\mathbb{R}$ is the real numbers.

10.2.12 Exercise (from the exam on 2002-10-21)**.** Decide for each of the following formulas whether it is a tautology or not. A complete explanation is required.

a) $\exists x_2 \forall x_1 (P_1(x_1) \leftrightarrow P_1(x_2))$

b) $\forall x_1 \exists x_2 (P_1(x_1) \leftrightarrow P_1(x_2))$

10.2.13 Exercise (from the exam on 2003-10-20)**.** Interpret the formula $\forall x_0 \forall x_1 (f_1(x_0) \doteq f_1(x_1) \rightarrow x_0 \doteq x_1)$ in the following structures and find its truth value in each one of them. Motivate!

a) $\langle \mathbb{N};\leq;s,0\rangle$, where s is the successor operation.

b) $\langle \mathbb{R};\leq;\sin,0\rangle$

10.3 Bounded quantifiers

In some occasions we would like to say "all" without talking about *all* elements, but rather *all those with a given property.* In the same way, we would like to say "some" in the sense of *some element with a given property.* For example: "Not all prime numbers are odd, some prime number is even". If the domain consists of the natural numbers and P_1 is interpreted as "is prime", while P_2 is interpreted as "is odd" and P_3 as "is even", we can express "all prime numbers are odd" as:

$$\forall x (P_1(x) \rightarrow P_2(x)) \tag{10.3.1}$$

and "some prime numbers are odd" as

$$\exists x(P_1(x) \wedge P_2(x)) . \tag{10.3.2}$$

In the same way we express "some prime number is even" as

$$\exists x(P_1(x) \wedge P_3(x)) . \tag{10.3.3}$$

More generally formulated: we express "all those with the property $P_1^{\mathcal{A}}$ have the property $P_2^{\mathcal{A}}$" as (10.3.1) and "some with the property $P_1^{\mathcal{A}}$ has the property $P_2^{\mathcal{A}}$" as (10.3.2).

10.3.4 Exercise (from the exam on 2005-08-23)**.** Formalize the proposition below; that is to say, give three formulas $\gamma_a, \gamma_b, \gamma_c$ that are interpreted as the three propositions (a, b, c) in the structure $\langle I; C, E, S; \rangle$, where I is the set of all curves in a plane, $C(x)$ is the predicate that aserts that x is a circle, $E(x)$ is the predicate that asserts that x is an ellipse, and $S(x, y)$ is the relation that asserts that x and y intersect. The formulas should not contain free variables.

a) All circles are ellipses.

b) Some ellipses are circles.

c) Every ellipse intersects some circle.

10.3.5 Exercise (from the exam on 2004-10-18)**.** In this exercise we use the arity type $\langle ; 2, 2, 0\rangle$.

a) Formalize the following propositions:

   No odd number is even

   More precisely: Give a formula φ such that its interpretation in $\langle \mathbb{N}; ; +, \cdot, 1\rangle$ is "for all odd x one has x is not even".

   To be odd is defined here as being equal to $2n+1$ for some natural number n.

   To be even is defined here as being equal to $2n$ for some natural number n.

b) Interpret the formula φ (from the previous item) in $\langle \mathbb{R}; ; +, \cdot, 1\rangle$ and give its truth value in this structure.

10.4 Summary

We have defined interpretation in predicate logic. An interpretation is given by a structure together with a valuation for the variables. Given an interpretation, every term gets a *value*, which is an element in the domain of the interpretation, and every formula gets a *truth value*, which is an element in $\{0, 1\}$, decided by the interpretation. If a formula has the truth value 1, one says that it is *true in the interpretation*, otherwise one says that it is *false in the interpretation.* One also says that a formula is interpreted as a proposition, which one gets by substituting $\wedge$ with *and*, $\vee$ with *or*, $\forall$ with *for all*, and so on. The linguistic ambiguities that can arise with such an interpretation are compensated by the fact that the truth value of a formula is precisely defined: for instance, it is clear that *or* has to be interpreted as *inclusive* from the fact that its truth value is computed in this way. In this chapter we have also introduced *reevaluations.* The most important thing to bring with you for the rest of this course is the ability to compute truth values of formulas in different interpretations and manipulate reevaluations, since these will be extensively used in many examples, exercises and proofs in what follows.

Chapter 11

Simplifications

11.1 Algebraic simplifications

In the same way as we do in propositional logic, we let $\approx$ mean that two formulas have always the same truth value.

11.1.1 Definition (logical equivalence)**.** By $\varphi \approx \psi$ (that φ and ψ are *logically equivalent*) we mean that $[\![\varphi]\!]^{\mathcal{A}} = [\![\psi]\!]^{\mathcal{A}}$ in all interpretations $\mathcal{A}$.

Since the truth value of formulas which are constructed by propositional operations has been defined precisely as in propositional logic, we can compute using Boolean algebra in predicate logic as well. However, we need new rules to compute with $\forall$ and $\exists$. They are collected in Figure 11.1, and we will now check that they are correct. Some of them are verified in the examples, and others are left for you as exercises.

$$\forall x(\varphi \wedge \psi) \approx \forall x\varphi \wedge \forall x\psi$$
$$\exists x(\varphi \vee \psi) \approx \exists x\varphi \vee \exists x\psi$$

$$\neg\forall x\varphi \approx \exists x\neg\varphi$$
$$\neg\exists x\varphi \approx \forall x\neg\varphi$$

$$\forall x\varphi \approx \varphi \qquad \text{if } x \text{ does not occur freely in } \varphi$$
$$\exists x\varphi \approx \varphi \qquad \text{if } x \text{ does not occur freely in } \varphi$$

$$\forall x(\varphi \vee \psi) \approx \forall x\varphi \vee \psi \qquad \text{if } x \text{ does not occur freely in } \psi$$
$$\exists x(\varphi \wedge \psi) \approx \exists x\varphi \wedge \psi \qquad \text{if } x \text{ does not occur freely in } \psi$$

Figure 11.1: Some useful computation rules in algebraic predicate logic

11.1.2 Example. Show that $\forall x\neg\varphi \approx \neg\exists x\varphi$.

Proof. Assume first that $[\![\forall x\neg\varphi]\!] = 1$. This means that $[\![\neg\varphi]\!]^{[x\mapsto a]} = 1$ for all a, which is the same as $[\![\varphi]\!]^{[x\mapsto a]} = 0$ for all a. But this says that $[\![\exists x\varphi]\!] = 0$, and hence $[\![\neg\exists x\varphi]\!] = \neg[\![\exists x\varphi]\!] = \neg 0 = 1$.

The argument can also be done backwards, proving that if $[\![\neg\exists x\varphi]\!] = 1$, then $[\![\forall x\neg\varphi]\!] = 1$. □

11.1.3 Exercise. Show that $\exists x\neg\varphi \approx \neg\forall x\varphi$.

11.1.4 Example. $\forall x(\varphi \wedge \psi) \approx \forall x\varphi \wedge \forall x\psi$.

Solution. Assume first that $[\![\forall x(\varphi \wedge \psi)]\!] = 1$. This means that $[\![\varphi \wedge \psi]\!]^{[x\mapsto a]} = 1$ for all a in the domain. It follows that $[\![\varphi]\!]^{[x\mapsto a]} = 1$ for all a in the domain, and similarly for ψ. Thus, it follows that $[\![\forall x\varphi]\!] = 1$ and $[\![\forall x\psi]\!] = 1$. Therefore, we have $[\![\forall x\varphi \wedge \forall x\psi]\!] = [\![\forall x\varphi]\!] \wedge [\![\forall x\psi]\!] = 1 \wedge 1 = 1$.

If, on the other hand, $[\![\forall x\varphi \wedge \forall x\psi]\!] = 1$ then we can, by following the previous argument backwards, assert that $[\![\forall x(\varphi \wedge \psi)]\!] = 1$. ☐

11.1.5 Example. $\exists x(\varphi \vee \psi) \approx \exists x\varphi \vee \exists x\psi$.

Solution. Assume that $[\![\exists x(\varphi \vee \psi)]\!] = 1$. This means that $[\![\varphi \vee \psi]\!]^{[x\mapsto a]} = 1$ for some a. Then we have either $[\![\varphi]\!]^{[x\mapsto a]} = 1$ or $[\![\psi]\!]^{[x\mapsto a]} = 1$. We consider the first case (the other one is completely analogous). We then have that $[\![\exists x\varphi]\!] = 1$, and it follows that $[\![\exists x\varphi \vee \exists x\psi]\!] = 1 \vee [\![\exists x\psi]\!] = 1$.

By following the argument backwards we can show the other direction of the equivalence. ☐

11.1.6 Example. Show that if x does not occur free in φ, then we have $\forall x\varphi \approx \varphi$.

Solution. That $\forall x\varphi \vDash \varphi$ holds follows from Exercise 10.2.2. To show the converse we assume that $[\![\varphi]\!] = 1$. We shall prove that $[\![\forall x\varphi]\!] = 1$, which by definition means we just need to check that $[\![\varphi]\!]^{[x\mapsto a]} = 1$ for all a in the domain. But according to Theorem 10.1.28 we have, since x does not occur free in φ, that $[\![\varphi]\!]^{[x\mapsto a]} = [\![\varphi]\!] = 1$. ☐

11.1.7 Exercise. Show that if x does not occur free in φ then we have $\exists x\varphi \approx \varphi$.

One cannot always *replace* a formula with another that has the same truth value.

There is a complication which makes algebraic simplifications in predicate logic not as slick as in propositional logic. If, for instance, φ is true, it would be tempting to substitute $\forall x_0\varphi$ with $\forall x_0\top$, but this is sometimes not correct: it can change the truth value. If, for example, $\varphi = \neg(x_0 \doteq x_1)$, then $\forall x_0\varphi$ is false (even when φ is true) while $\forall x_0\top$ is true. To replace a formula by another, it is not sufficient that they have the same truth value in the interpretation we are working with, but the truth values need to be the same in *all* interpretations. Theorem 11.1.8 shows that it is sufficient.

11.1.8 Theorem. *If $\varphi \approx \psi$, then φ can be replaced with ψ in any formula without changing its truth value.*

Proof. Consider the definition of truth value (10.1.19). It is given by recursion, so that the truth value is given by the truth value of the subformulas. Therefore, if one replaces one subformula with another of the same truth value, the result will not be affected. The only difficulty appears with the case of quantifiers, where we do not use the same valuation for subformulas but use reevaluations instead. Since we assume that φ and ψ have te same value in *every* valuation, they are guaranteed to have the same value even in the reevaluations. Hence, the result follows. ☐

In the case $[\![\psi]\!] = 0$ it would have been tempting to replace $\forall x(\varphi \vee \psi)$ with $\forall x(\varphi \vee \bot)$, but as we have shown above, such replacements sometimes give wrong results. We must use that x does not occur freely in ψ.

We now continue giving examples of logically equivalent formulas. The following proof illustrates how one sometimes does not have enough hypothesis to use the previous theorem.

11.1.9 Example. $\forall x(\varphi \vee \psi) \approx \forall x\varphi \vee \psi$ if x does not occur freely in ψ (compare examples 10.2.7, 10.2.8).

Solution. We consider two cases: when $[\![\psi]\!] = 0$, respectively $[\![\psi]\!] = 1$.

In the first case we have $[\![\forall x\varphi \vee \psi]\!] = [\![\forall x\varphi]\!]$. We shall prove that $[\![\forall x(\varphi \vee \psi)]\!] = [\![\forall x\varphi]\!]$. Assume therefore that $[\![\forall x(\varphi \vee \psi)]\!] = 1$. This means that $[\![\varphi \vee \psi]\!]^{[x\mapsto a]} = 1$ for all a. Hence we have, for every a, either $[\![\varphi]\!]^{[x\mapsto a]} = 1$ or

$[\![\psi]\!]^{[x\mapsto a]} = 1$; but the latter is impossible, since $[\![\psi]\!]^{[x\mapsto a]} = [\![\psi]\!]$ according to Theorem 10.1.28 and the fact that $[\![\psi]\!] = 0$ by assumption. We have, hence, $[\![\varphi]\!]^{[x\mapsto a]} = 1$ for every a, and thus, $[\![\forall x\varphi]\!] = 1$. Assume, on the other hand that $[\![\forall x\varphi]\!] = 1$. This means that $[\![\varphi]\!]^{[x\mapsto a]} = 1$ for all a. Thus, $[\![\varphi \vee \psi]\!]^{[x\mapsto a]} = 1$ for all a, which gives $[\![\forall x(\varphi \vee \psi)]\!] = 1$.

In the case when $[\![\psi]\!] = 1$ we have to show that both sides have truth value 1. This follows by an argument which ressembles the one above but is simpler. You can probably do it yourself, if you have understood the proof so far. □

11.1.10 Exercise. Show that $\exists x(\varphi \wedge \psi) \approx \exists x\varphi \wedge \psi$ if x does not occur freely in ψ.

11.1.11 Exercise (from the exam on 2004-01-08)**.** Show that $\forall x(\varphi \vee \psi) \approx \forall x\varphi \vee \forall x\psi$ is not true in general. Show also that it does hold if x does not occur freely in ψ.

Let this exercise be a warning for sloppy simplifications!

11.1.12 Exercise (from the exam on 2004-08-17)**.** Give a (preferably natural) example from mathematics where the difference between $\forall x_0\exists x_1\varphi$ and $\exists x_1\forall x_0\varphi$ is exhibited.

11.1.13 Exercise. Decide whether $\exists x_0(P_1(x_0) \to \forall x_0 P_1(x_0))$ is true in all interpretations.

11.1.14 Exercise (from the exam on 2004-10-18)**.** Decide whether $(\forall x_0 P_1(x_0) \to \forall x_0 P_2(x_0)) \to \forall x_0(P_1(x_0) \to P_2(x_0))$ is true in all interpretations.

11.1.15 Exercise. Show by using algebra that, for all formulas φ, ψ, it is true in general that

a) $\neg\forall x(\varphi \to \psi) \approx \exists x(\varphi \wedge \neg\psi)$

b) $\neg\exists x(\varphi \wedge \psi) \approx \forall x(\varphi \to \neg\psi)$

Here one sees some of the duality between $\wedge$ and $\to$ described by the Galois connection (2.2.3).

11.2 Simplification by substitution

This is used in the proof of the soundness theorem (Chapter 13).

It is important to have a complete understanding of how the interpretation of a term or a formula is affected by substitution. The following theorem clarifies the situation in the case of terms. It says that, through substitution, the value is changed in the same way as if we replaced the value of the variable by the value of the inserted term.

11.2.1 Theorem. $[\![s[t/x_j]]\!] = [\![s]\!]^{[x_j\mapsto [\![t]\!]]}$

Proof. As usual, we give a proof by induction. If $s = x_i$, we have

$$
\begin{aligned}
[\![s[t/x_j]]\!] &= [\![x_i[t/x_j]]\!] &(11.2.2)\\
&= \begin{cases} [\![t]\!] & \text{if } j = i \\ [\![x_i]\!] & \text{otherwise} \end{cases} &(11.2.3)\\
&= [\![x_i]\!]^{[x_j\mapsto [\![t]\!]]} &(11.2.4)\\
&= [\![s]\!]^{[x_j\mapsto [\![t]\!]]}. &(11.2.5)
\end{aligned}
$$

If $s = f_i(t_1, \dots, t_{a_i})$, we have

$$
\begin{aligned}
[\![s[t/x_j]]\!] &= [\![f_i(t_1, \dots, t_{a_i})[t/x_j]]\!] &(11.2.6)\\
&= [\![f_i(t_1[t/x_j], \dots, t_{a_i}[t/x_j])]\!] &(11.2.7)\\
&= f_i^{\mathcal{A}}([\![t_1[t/x_j]]\!], \dots, [\![t_{a_i}[t/x_j]]\!]) &(11.2.8)\\
&= f_i^{\mathcal{A}}([\![t_1]\!]^{[x_j\mapsto [\![t]\!]]}, \dots, [\![t_{a_i}]\!]^{[x_j\mapsto [\![t]\!]]}) &(11.2.9)\\
&= [\![f_i(t_1, \dots, t_{a_i})]\!]^{[x_j\mapsto [\![t]\!]]} &(11.2.10)\\
&= [\![s]\!]^{[x_j\mapsto [\![t]\!]]}. &(11.2.11)
\end{aligned}
$$

□

It would have been good to have a similar theorem for the case when the term s is replaced by a formula φ. Unfortunately, we do not have such a theorem in general, but we have to add one more hypothesis concerning the substitution, namely, that *t is free for x in φ*. This notion requires some motivation.

Consider, for example, the formula $\forall x_0 \exists x_1 P_1(x_0, x_1)$, where we let the domain be the real numbers and interpret P_1 as the relation $<$. It says that for every real number there is a greater real number. Since x_0 is interpreted as a real number, then $\exists x_1 P_1(x_0, x_1)$ is true in the interpretation, no matter which value $[\![x_0]\!]$ has. But any term t can be interpreted as a real number, so one might think that

$$(\exists x_1 P_1(x_0, x_1))[t/x_0] \qquad (11.2.12)$$

should always be true. However, it is not true if $t = x_1$.

11.2.13 Exercise. Simplify (11.2.12) and compute its truth value when $t = x_1$.

Therefore, we cannot expect a theorem saying that

$$[\![\varphi[t/x_j]]\!] = [\![\varphi]\!]^{[x_j \mapsto [\![t]\!]]} \qquad (11.2.14)$$

holds in general, since there are exceptions.

11.2.15 Exercise. Give examples of φ, t and j such that (11.2.14) is false in the interpretation we have used in the previous exercise and example.

You can compare the above with the claim that

$$\int_0^1 xy \, \mathrm{d}x = y/2 \qquad (11.2.16)$$

holds for all y. When one says such a thing, one does not think that one could let x be y and conclude that

$$\int_0^1 x^2 \, \mathrm{d}x = x/2\,. \qquad (11.2.17)$$

This is however often done by high school students, and it is not so strange: no one has told them that one has to be careful when doing substitutions, and the values will only be reasonable if the inserted term is *free for* the variable one inserted it for, which means that no variable in the term is bound by any quantifier when doing the substitution. In ordinary mathematics, we avoid completely substitution of terms which are not free for the variables we substitute, but we usually forget to teach that this is the case. In logic we specify instead which substitutions give sensible results by defining *free for* formally. We make it simpler by first defining *bound for*. The idea behind the definition is that a term t is bound by a variable x in a formula if the substitution $[t/x]$ leads to some variable in t being bound by a quantifier. We state the definition by recursion.

11.2.18 Definition. t is bound for x_i in ...

$$
\begin{aligned}
\dots\ (t_1 \doteq t_2) &\overset{\text{def}}{=} \text{false}\\
\dots\ P_i(t_1,\dots,t_{r_i}) &\overset{\text{def}}{=} \text{false}\\
\dots\ \top &\overset{\text{def}}{=} \text{false}\\
\dots\ \bot &\overset{\text{def}}{=} \text{false}\\
\dots\ \varphi\wedge\psi &\overset{\text{def}}{=} t \text{ is bound for } x_i \text{ in at least one of } \varphi,\psi\\
\dots\ \varphi\vee\psi &\overset{\text{def}}{=} t \text{ is bound for } x_i \text{ in at least one of } \varphi,\psi\\
\dots\ \varphi\to\psi &\overset{\text{def}}{=} t \text{ is bound for } x_i \text{ in at least one of } \varphi,\psi\\
\dots\ \forall x_j\varphi &\overset{\text{def}}{=} \bullet\ i\neq j \text{ and}\\
&\quad\ \bullet\ x_i \text{ occurs freely in } \varphi \text{ and}\\
&\quad\ \bullet\ x_j \text{ occurs in } t, \text{ or } t \text{ bound for } x_i \text{ in } \varphi\\
\dots\ \exists x_j\varphi &\overset{\text{def}}{=} \bullet\ i\neq j \text{ and}\\
&\quad\ \bullet\ x_i \text{ occurs freely in } \varphi \text{ and}\\
&\quad\ \bullet\ x_j \text{ occurs in } t, \text{ or } t \text{ bound for } x_i \text{ in } \varphi
\end{aligned}
$$

11.2.19 Definition. t is free for x_i in φ if t is not bound for x_i in φ.

Did you understand the difference between *free in* and *free for*?

11.2.20 Exercise.

a) Show that x_1 is free for x_0 in $\exists x_0 P_1(x_0, x_1)$.

b) Show that x_0 is bound for x_1 in the same formula.

c) Is x_0 free for x_1 in $\forall x_0 P_1(x_0)$?

11.2.21 Exercise (from the exam on 2003-01-09)**.** In which of the substitutions in Exercise 9.2.19 b are the inserted terms free for the variables one substitutes?

11.2.22 Exercise. Show that x is free for x in φ.

11.2.23 Exercise. Show that if t is bound for x in φ, then some of the variables in t are quantified in φ.

11.2.24 Exercise. Show that if none of the variables in t are quantified in φ, then t is free for x in φ.

This gives in many cases an efficient way to see if we have the *free for*-property.

The condition in the following theorem is the reason why the notion *free for* is so important. The theorem says that under this condition, substitution works the way we would like regarding the truth values.

11.2.25 Theorem. *If t is free for x in φ then we have $[\![\varphi[t/x]]\!] = [\![\varphi]\!]^{[x\mapsto[\![t]\!]]}$.*

Proof. We give a proof by induction (induction on the structure of Form), and therefore we go through all different forms that formulas can have.

If φ has some of the forms $t_1 \doteq t_2$, $P_i(t_1,\dots,t_{r_i})$, $\top$ or $\bot$, then it is always true that t is free for x in φ according to the definition of *free for* and *bound for*. In this case it is also easy to check that it holds by applying Theorem 11.2.1.

If φ is composed by using $\wedge$, $\vee$ or $\to$, the theorem follows immediately from the inductive hypothesis.

We will consider the case when φ is of the form $\forall x_j\psi$. Say that $x = x_i$; then we will show that

$$[\![(\forall x_j\psi)[t/x_i]]\!] = [\![\forall x_j\psi]\!]^{[x_i\mapsto[\![t]\!]]}\,. \tag{11.2.26}$$

Note first that if x_i does not occur freely in $\forall x_j\psi$, then both sides of (11.2.26) are simplified to $[\![\forall x_j\psi]\!]$, from which the result follows immediately. We therefore assume in what follows that x_i occurs freely in $\forall x_j\psi$; that is, that $i \neq j$ and x_i occurs freely in ψ. The assumption that t is free for x_i in $\forall x_j\psi$ means that t is not bound for x_i, and since $i\neq j$ and x_i occurs freely in ψ, the following must be false:

- x_j occurs in t, or t is bound for x_i in ψ.

We draw the conclusion that

- x_j does not occur in t,
- t is free for x_i in ψ.

The way of using the inductive hypothesis has similarities with the way it is used in the proof of Theorem 10.1.28.

The left hand side of (11.2.26) may now, since $i \neq j$, be simplified to $[\![\forall x_j \psi[t/x_i]]\!]$. Assume that its value is 1. This means that $[\![\psi[t/x_i]]\!]^{[x_j \mapsto a]} = 1$ for all a in the domain. Because of the inductive hypothesis, we have that

$$[\![\psi]\!]^{[x_j \mapsto a][x_i \mapsto [\![t]\!]^{[x_j \mapsto a]}]} = 1\,. \tag{11.2.27}$$

Since x_j does not occur in t, we can simplify this to

$$[\![\psi]\!]^{[x_j \mapsto a][x_i \mapsto [\![t]\!]]} = 1 \tag{11.2.28}$$

and since $i \neq j$, we can change the ordering (Exercise 10.1.18):

$$[\![\psi]\!]^{[x_i \mapsto [\![t]\!]][x_j \mapsto a]} = 1\,. \tag{11.2.29}$$

But this means precisely that the right side of (11.2.26) is 1. By following this reasoning backwards, the other direction of the equivalence is shown.

The case of $\exists$ formulas is completely analogous. □

11.2.30 Example. Show that $\varphi[t/x] \vDash \exists x \varphi$ if t is free for x in φ.

11.2.30 and 11.2.31 will be later used in the proof of the soundness theorem (Theorem 13.1.1).

Solution. Assume that t is free for x in φ and that $[\![\varphi[t/x]]\!] = 1$. We will prove that $[\![\exists x \varphi]\!] = 1$. That $[\![\varphi[t/x]]\!] = 1$ gives, according to the previous theorem, that $[\![\varphi]\!]^{[x \mapsto [\![t]\!]]} = 1$, but then we can take $a = [\![t]\!]$, so we have that $[\![\varphi]\!]^{[x \mapsto a]} = 1$, which means that $[\![\exists x \varphi]\!] = 1$. □

11.2.31 Exercise. Show that $\forall x \varphi \vDash \varphi[t/x]$ if t is free of x in φ.

11.2.32 Exercise. Show that $\varphi[t/x] \vDash \exists x \varphi$ does not necessarily hold if t is bound for x in φ.

Hint. Let $x = x_0$, $t = x_1$, $\varphi = \forall x_1 (x_0 \doteq x_1)$.

11.2.33 Exercise. Show that $\forall x \varphi \vDash \varphi[t/x]$ does not necessarily hold if t is bound for x in φ.

This example illustrates the difference between substitution and reevaluation: by the simplification of $t[f_2/x_i][f_1/x_i]$, it is the second square bracket which is deleted, while by the simplification of $\mathcal{A}[x_i \mapsto a][x_i \mapsto b]$ it is the first square bracket the one that is deleted. The difference is explained by the fact that substitution changes terms, while reevaluation changes valuations.

11.2.34 Example (cf. Exercise 10.1.17 a)**.** Let $f_1^{\mathcal{A}} = a$ and $f_2^{\mathcal{A}} = b$. Simplify $[\![t[f_2/x_i][f_1/x_i]]\!]$.

Solution. We can do this in two ways. On one hand we can use that x_i does not occur in $t[f_2/x_i]$ (which is shown by an inductive proof) and therefore conclude that $t[f_2/x_i][f_1/x_i] = t[f_2/x_i]$ (Exercise 9.1.18), so that

$$[\![t[f_2/x_i][f_1/x_i]]\!] = [\![t[f_2/x_i]]\!] = [\![t]\!]^{[x_i \mapsto b]}\,. \tag{11.2.35}$$

On the other hand we can compute, with the help of Theorem 11.2.1,

$$[\![t[f_2/x_i][f_1/x_i]]\!] = [\![t[f_2/x_i]]\!]^{[x_i \mapsto a]} = [\![t]\!]^{[x_i \mapsto a][x_i \mapsto b]} \tag{11.2.36}$$

and use Exercise 10.1.17 a to conclude that this is $[\![t]\!]^{[x_i \mapsto b]}$. □

11.2.37 Exercise (cf. Exercise 9.2.22)**.** Simplify

a) $[\![t[y/x][x/y]]\!]$

b) $[\![\varphi[y/x][x/y]]\!]$ if x is free for y in $\varphi[y/x]$ and y is free for x in φ.

11.2.38 Exercise (this is used in the proof of the completeness theorem, Chapter 14)**.** Show through an inductive proof that if y does not occur in ψ, then x is free for y in $\psi[y/x]$.

> *Hint.* Here is a sketch of the proof; do the details by yourself. Use induction on the complexity of the formula. The induction step is easy in the case when ψ is of one of the forms $\forall x_j \varphi$ and $\exists x_j \varphi$. Consider one of the cases, the other one is completely analogous. Assume that $\psi = \forall x_j \varphi$ and that y does not occur in ψ. We shall show that x is free for y in $\psi[y/x]$, and the inductive hypothesis we can use is that x is free for y in $\varphi[y/x]$. Consider two cases. If $x = x_j$, then $\psi[y/x] = \psi$, in which case y does not occur at all, so we are done. If $x \neq x_j$, then we have that $\psi[y/x] = \forall x_j \varphi[y/x]$. But here x is free for y since x_j does not occur in x (which is clear by the definition of "occurs in"), and since x is free for y in $\varphi[y/x]$ (inductive hypothesis).

11.3 Summary

We have seen how with the help of algebraic simplifications and substitution simplification we can compute the truth value of formulas in a considerably easier way. We have also seen that the notion *free for* is very important in this context: simplification by substitution is not guaranteed to work when terms are bound for variables. The most important thing to take with you for the rest of this course is the skill to simplify the computation of truth values by algebraic methods, as well as the ability to use theorems 11.2.1 and 11.2.25. This includes understanding what *free for* means and how we can decide whether a term is free for a variable in a formula; otherwise, you would not be able to use the theorems in the right way.

Chapter 12

Natural deduction

12.1 New rules

Natural deduction in predicate logic is done precisely in the same way as in propositional logic, but with even more rules. These are collected in Figure 12.1 (page 90). The rule “refl” is called *reflexivity* and the rule “repl” is called the *replacement rule*.

Note the various restrictions appearing in the rules. To be able to use some of them, it is required that some terms are free for some variables, while for some other rules it is required that variables do not occur freely in certain formulas. These restriction are important – disregarding them can lead to deriving false formulas.

The principles for derivations are otherwise the same as in propositional natural deduction. This chapter, therefore, does not contain any theory; only examples and exercises. We just have to modify some definitions.

▶ **12.1.1 Definition.** By $\varphi_1, \dots, \varphi_n \vdash \varphi$ we mean that there is a derivation of φ, with only the rules of figure 5.1 and 12.1 and without any undischarged assumptions, except possibly $\varphi_1, \dots, \varphi_n$. (Compare Definition 5.5.1.)

12.1.2 Example. Show that $\vdash (x_0 \doteq x_0)$.

Solution. Since x_0 is a term, we can use the rule for reflexivity.

$$\dfrac{}{x_0 \doteq x_0}\,\text{refl}$$

□

12.1.3 Example (symmetry)**.** Show that $t \doteq s \vdash s \doteq t$.

Solution. If we let $\varphi = (x \doteq t)$, where we choose x so that it does not occur in t, we get

$$\varphi[t/x] = (x[t/x] \doteq t[t/x]) = (t \doteq t)\,, \tag{12.1.4}$$
$$\varphi[s/x] = (x[s/x] \doteq t[s/x]) = (s \doteq t)\,, \tag{12.1.5}$$

and we can use the rule for replacement.

$$\dfrac{\dfrac{}{t \doteq t}\,\text{refl} \qquad t \doteq s}{s \doteq t}\,\text{ers}$$

□

Note that the formula φ does not occur in the derivation. In fact, we cannot see which variable x has been chosen and therefore we cannot see which formula φ was in the replacement rule. However, we can always decide whether an application of the replacement rule is correct by noting how the formulas above and under the line differ. With that information, one can see whether there exists a formula φ which can be used in the rule, but the choice is not unique.

12.1.6 Exercise (transitivity)**.** Show that $u \doteq t, t \doteq s \vdash u \doteq s$.

Hint. Find the formula φ which could be used together with the replacement rule.

φ, σ denote arbitrary formulas
t, s denote arbitrary terms
x denote an arbitrary variable

When substituting, it is assumed that t (resp. s) are free for x.

$$\frac{}{t \doteq t}\ \text{refl} \qquad\qquad \frac{\varphi[t/x] \qquad t \doteq s}{\varphi[s/x]}\ \text{ers}$$

$$\frac{\varphi}{\forall x \varphi}\ \forall I \qquad\qquad \frac{\forall x \varphi}{\varphi[t/x]}\ \forall E$$

($\forall I$) where x does not occur freely in some undischarged assumption.

$$\frac{\varphi[t/x]}{\exists x \varphi}\ \exists I \qquad\qquad \frac{\exists x \varphi \qquad \begin{matrix}[\varphi] \\ \vdots \\ \sigma\end{matrix}}{\sigma}\ \exists E$$

($\exists E$) where x does not occur freely in σ, nor in any undischarged assumption in the right subtree, except possibly in φ.

Figure 12.1: Additional rules for natural deduction in predicate logic

12.1.7 Example. Construct a derivation of $\forall x_0 \forall x_1 (x_0 \doteq x_1 \to x_1 \doteq x_0)$.

Solution.

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{x_0 \doteq x_0}\,\text{refl} \qquad [x_0 \doteq x_1]^1}{x_1 \doteq x_0}\,\text{ers}}{x_0 \doteq x_1 \to x_1 \doteq x_0}\,{\to}I_1}{\forall x_1 (x_0 \doteq x_1 \to x_1 \doteq x_0)}\,\forall I}{\forall x_0 \forall x_1 (x_0 \doteq x_1 \to x_1 \doteq x_0)}\,\forall I$$

Here we have an implication introduction which discharges the assumption $x_0 \doteq x_1$. This is why the $\forall$-introductions are allowed. □

Do you see that both uses of $\forall I$ would be forbidden if the assumption $x_0 \doteq x_1$ was undischarged? Otherwise, check the rules in Figure 12.1.

12.1.8 Exercise. Construct a derivation of

$$\forall x_0 \forall x_1 \forall x_2 ((x_0 \doteq x_1) \land (x_1 \doteq x_2) \to (x_0 \doteq x_2)) \, .$$

A special case of $\forall E$ is

$$\dfrac{\forall x \varphi}{\varphi}\,\forall E \tag{12.1.9}$$

which one gets by putting $t = x$, since $\varphi[x/x] = \varphi$ (Exercise 9.2.21). It is the most common way to use the rule. It occurs, for example, in the solution of the following problem.

12.1.10 Example. Show that $\forall x \neg \varphi \vdash \neg \exists x \varphi$.

Solution.

$$\dfrac{\dfrac{[\exists x \varphi]^2 \qquad \dfrac{\dfrac{\forall x \neg \varphi}{\neg \varphi}\,\forall E \qquad [\varphi]^1}{\bot}\,{\to}E}{\bot}\,\exists E_1}{\neg \exists x \varphi}\,{\to}I_2$$

We must check that the variable restrictions are satisfied. The only rule we use which has such restrictions is $\exists E$. It requires, in the above case, that x does not occur freely in $\bot$ (which is not the case), and that x does not occur freely in some undischarged assumption in the derivation of $\bot$, except possibly in φ. In our case, $\forall x \neg \varphi$ is the only undischarged assumption, except from φ, and x does not occur freely in it (it occurs bound, however, but that is not a problem).

On the exam you do not need to justify why the variable restrictions are satisfied if you are not explicitely instructed to do so. Otherwise, a derivation is considered to be *wrong* if the restrictions are not satisfied.

In what follows, the fact that the variable restrictions need to be satisfied will not be explicitly checked, but these checkings must always be done before one can assert that the derivation is correct. □

12.1.11 Example. Show that $\neg \exists x \varphi \vdash \forall x \neg \varphi$.

Solution.

$$\dfrac{\dfrac{\dfrac{\neg \exists x \varphi \qquad \dfrac{[\varphi]^1}{\exists x \varphi}\,\exists I}{\bot}\,{\to}E}{\neg \varphi}\,{\to}I_1}{\forall x \neg \varphi}\,\forall I$$

□

12.1.12 Example. Show that $\exists x \neg \varphi \vdash \neg \forall x \varphi$.

Solution.

$$\dfrac{\dfrac{\exists x\neg\varphi \qquad \dfrac{[\neg\varphi]^1 \qquad \dfrac{[\forall x\varphi]^2}{\varphi}\,\forall E}{\bot}\,{\to}E}{\bot}\,\exists E_1}{\dfrac{\bot}{\neg\forall x\varphi}\,{\to}I_2}$$

□

12.1.13 Example. Show that $\neg\forall x\varphi \vdash \exists x\neg\varphi$.

Solution. Here we must use RAA twice.

$$\dfrac{\dfrac{\neg\forall x\varphi \qquad \dfrac{\dfrac{\dfrac{[\neg\exists x\neg\varphi]^2 \qquad \dfrac{[\neg\varphi]^1}{\exists x\neg\varphi}\,\exists I}{\bot}\,{\to}E}{\varphi}\,\mathrm{RAA}_1}{\forall x\varphi}\,\forall I}{\bot}\,{\to}E}{\exists x\neg\varphi}\,\mathrm{RAA}_2$$

□

12.1.14 Example. Show that $\vdash \forall x(\varphi\wedge\psi) \leftrightarrow \forall x\varphi\wedge\forall x\psi$.

Solution.

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{[\forall x(\varphi\wedge\psi)]^1}{\varphi\wedge\psi}\,\forall E}{\varphi}\,\wedge E}{\forall x\varphi}\,\forall I \qquad \dfrac{\dfrac{\dfrac{[\forall x(\varphi\wedge\psi)]^1}{\varphi\wedge\psi}\,\forall E}{\psi}\,\wedge E}{\forall x\psi}\,\forall I}{\forall x\varphi\wedge\forall x\psi}\,\wedge I}{\forall x(\varphi\wedge\psi)\to\forall x\varphi\wedge\forall x\psi}\,{\to}I_1 \qquad \dfrac{\dfrac{\dfrac{\dfrac{\dfrac{[\forall x\varphi\wedge\forall x\psi]^2}{\forall x\varphi}\,\wedge E}{\varphi}\,\forall E \qquad \dfrac{\dfrac{[\forall x\varphi\wedge\forall x\psi]^2}{\forall x\psi}\,\wedge E}{\psi}\,\forall E}{\varphi\wedge\psi}\,\wedge I}{\forall x(\varphi\wedge\psi)}\,\forall I}{\forall x\varphi\wedge\forall x\psi\to\forall x(\varphi\wedge\psi)}\,{\to}I_2}{\forall x(\varphi\wedge\psi)\leftrightarrow\forall x\varphi\wedge\forall x\psi}\,\wedge I$$

□

The examples 12.1.14 and 12.1.15 show how the rules for $\forall$ and $\exists$ are used in more complicated cases. Note that, in Example 12.1.15, it is important to use $\exists E$ sufficiently far down in the derivation so that the variable restrictions are satisfied.

12.1.15 Example. Show that $\vdash \exists x(\varphi\vee\psi) \leftrightarrow \exists x\varphi\vee\exists x\psi$.

You can try to construct a derivation by yourself.

Solution. See Figure 12.2 (page 93). □

12.1.16 Example. a) What is wrong with the following derivation?

$$\dfrac{\dfrac{[\exists x_0(x_0\doteq x_1)]^1 \qquad \dfrac{[x_0\doteq x_1]^2}{\exists x_1(x_1\doteq x_1)}\,\exists I}{\exists x_1(x_1\doteq x_1)}\,\exists E_2}{\exists x_0(x_0\doteq x_1)\to\exists x_1(x_1\doteq x_1)\,.}\,{\to}I_1$$

b) Can one derive $\exists x_0(x_0\doteq x_1)\to\exists x_1(x_1\doteq x_1)$?

Solution. a) By $\exists I$ the formula under the line is $\exists x_1\varphi$, where $\varphi = (x_1\doteq x_1)$. The formula above the line should be of the form $\varphi[t/x_1]$; that is, $t\doteq t$ for some term t. But since x_0 and x_1 are different variables, this is not correct.

b) Sure, for instance:

$$\dfrac{\dfrac{\dfrac{}{x_1\doteq x_1}\,\mathrm{refl}}{\exists x_1(x_1\doteq x_1)}\,\exists I}{\exists x_0(x_0\doteq x_1)\to\exists x_1(x_1\doteq x_1)\,.}\,{\to}I$$

$$\dfrac{\dfrac{[\exists x(\varphi\lor\psi)]^3 \qquad \dfrac{[\varphi\lor\psi]^2 \qquad \dfrac{\dfrac{[\varphi]^1}{\exists x\varphi}\,\exists I}{\exists x\varphi\lor\exists x\psi}\,\lor I \qquad \dfrac{\dfrac{[\psi]^1}{\exists x\psi}\,\exists I}{\exists x\varphi\lor\exists x\psi}\,\lor I}{\exists x\varphi\lor\exists x\psi}\,\lor E_1}{\dfrac{\exists x\varphi\lor\exists x\psi}{\exists x(\varphi\lor\psi)\to\exists x\varphi\lor\exists x\psi}\,\to I_3}\,\exists E_2 \qquad \dfrac{[\exists x\varphi\lor\exists x\psi]^7 \qquad \dfrac{[\exists x\varphi]^6 \qquad \dfrac{\dfrac{[\varphi]^4}{\varphi\lor\psi}\,\lor I}{\exists x(\varphi\lor\psi)}\,\exists I}{\exists x(\varphi\lor\psi)}\,\exists E_4 \qquad \dfrac{[\exists x\psi]^6 \qquad \dfrac{\dfrac{[\psi]^5}{\varphi\lor\psi}\,\lor I}{\exists x(\varphi\lor\psi)}\,\exists I}{\exists x(\varphi\lor\psi)}\,\exists E_5}{\dfrac{\exists x(\varphi\lor\psi)}{\exists x\varphi\lor\exists x\psi\to\exists x(\varphi\lor\psi)}\,\to I_7}\,\lor E_6}{\exists x(\varphi\lor\psi)\leftrightarrow\exists x\varphi\lor\exists x\psi}\,\land I$$

Figure 12.2: Solution of Example 12.1.15.

Remember that one does not *have to* discharge anything when introducing implication. □

12.2 Misc. exercises

12.2.1 Exercise. Show that $\neg\exists x\varphi \vdash \neg\forall x\varphi$.

12.2.2 Exercise. Show that $\vdash \forall x\varphi \leftrightarrow \varphi$ if x does not occur freely in φ.

12.2.3 Exercise. Show that $\vdash \exists x\varphi \leftrightarrow \varphi$ if x does not occur freely in φ.

12.2.4 Exercise. Show that $\vdash \forall x(\varphi \vee \psi) \leftrightarrow \forall x\varphi \vee \psi$ if x does not occur freely in ψ. Notice in which part of the derivation this assumption is used.

12.2.5 Exercise. Show that $\vdash \exists x(\varphi \wedge \psi) \leftrightarrow \exists x\varphi \wedge \psi$ if x does not occur freely in ψ. Notice in which part of the derivation this assumption is used.

12.2.6 Exercise (from the exam on 2005-08-23)**.**

a) Construct a derivation of $(\exists x\varphi \to \psi) \to (\forall x\varphi \to \psi)$.

b) Construct a derivation of $(\forall x\varphi \to \psi) \to (\exists x\varphi \to \psi)$ that is correct if x does not occur freely in φ.

c) An attempt to derive $(\forall x\varphi \to \psi) \to (\exists x\varphi \to \psi)$ could be the following tree, but if x occurs freely in φ or ψ, the derivation is not correct. Explain what the error is and what is wrong. Point out the precise location of errors!

$$\dfrac{\dfrac{\dfrac{[\exists x\varphi]^2 \qquad \dfrac{\dfrac{[\varphi]^1}{\forall x\varphi}\,\forall I \qquad [\forall x\varphi \to \psi]^3}{\psi}\,{\to}E}{\psi}\,\exists E_1}{\exists x\varphi \to \psi}\,{\to}I_2}{(\forall x\varphi \to \psi) \to (\exists x\varphi \to \psi)}\,{\to}I_3$$

d) Show that, if $\psi = (x \doteq x)$, there is a correct derivation of $(\forall x\varphi \to \psi) \to (\exists x\varphi \to \psi)$.

12.2.7 Exercise (from the exam on 2003-01-09)**.** Derive $(\exists x P_1(x) \to \forall x P_2(x)) \leftrightarrow \forall x(\exists x P_1(x) \to P_2(x))$.

12.2.8 Exercise (from the exam on 2004-08-17)**.** Derive $\forall x\varphi \vee \exists x\neg\varphi$.

Hint. One has to use RAA several times.

12.2.9 Exercise (from the exam on 2004-08-17)**.** Explain why the following derivation is not correct if x_0 occurs freely in φ (specify precisely which step in the derivation is wrong and explain why).

$$\dfrac{\dfrac{\dfrac{[\forall x_0\exists x_1\varphi]^2}{\exists x_1\varphi}\,\forall E \qquad \dfrac{\dfrac{[\varphi]^1}{\forall x_0\varphi}\,\forall I}{\exists x_1\forall x_0\varphi}\,\exists I}{\exists x_1\forall x_0\varphi}\,\exists E_1}{\forall x_0\exists x_1\varphi \to \exists x_1\forall x_0\varphi}\,{\to}I_2$$

12.2.10 Exercise (from the exam on 2005-01-07)**.** Derive $\forall x(\neg\varphi \vee \neg\psi) \leftrightarrow \neg\exists x(\varphi \wedge \psi)$.

12.2.11 Exercise (from the exam on 2004-10-18)**.**

a) Explain why the following is not a correct derivation if $\psi = (x_0 \doteq x_0)$.

$$\dfrac{\dfrac{[\exists x_0(\varphi \wedge \psi)]^2 \qquad \dfrac{\dfrac{\dfrac{[\varphi \wedge \psi]^1}{\varphi}\,{\scriptstyle \wedge E}}{\exists x_0 \varphi}\,{\scriptstyle \exists I} \qquad \dfrac{[\varphi \wedge \psi]^1}{\psi}\,{\scriptstyle \wedge E}}{(\exists x_0 \varphi) \wedge \psi}\,{\scriptstyle \wedge I}}{(\exists x_0 \varphi) \wedge \psi}\,{\scriptstyle \exists E_1}}{\exists x_0(\varphi \wedge \psi) \to (\exists x_0 \varphi) \wedge \psi}\,{\scriptstyle \to I_2}$$

Specify precisely which step is wrong and explain why!

b) Show that there is a correct derivation of $\exists x_0(\varphi \wedge \psi) \to (\exists x_0 \varphi) \wedge \psi$ if $\psi = (x_0 \doteq x_0)$.

12.2.12 Exercise (from the exam on 2002-08-20)**.** Derive $\exists x\varphi \vee \psi \leftrightarrow \exists x(\varphi \vee \psi)$, where x does not occur freely in ψ.

12.2.13 Exercise (from the exam on 2002-10-21)**.** Derive $(\exists x P_1(x) \to \forall x P_2(x)) \leftrightarrow \forall x(P_1(x) \to \forall x P_2(x))$.

12.2.14 Exercise (from the exam on 2003-08-19)**.** Derive $(\psi \to \exists x\varphi) \leftrightarrow \exists x(\psi \to \varphi)$, where x does not occur freely in ψ.

12.2.15 Exercise (from the exam on 2003-10-20)**.** Derive $(\exists x\varphi \to \psi) \leftrightarrow \forall x(\varphi \to \psi)$, where x does not occur freely in ψ. Specify in which part of the derivation these assumptions are used.

12.3 Summary

We have extended the formal system with new rules to cover the new ingredients in the language. The rules from propositional logic still hold. The most important thing to remember from here is the ability to construct derivations by using both the old and the new rules. You should also be able to decide if yours or someone elses's derivation is correct, for which you need to know both the rules and the limitations that there are for the variables. For instance, one rule (which?) is only allowed to be used when a certain variable does not occur freely in any undischarged assumption, and another (which?) has a more complicated set of limitations. Remember also that every rule that contains a substitution in its formulation requires that the term is free for the variable in the formula.

Chapter 13

Soundness & Review exercises

13.1 Soundness

We have already put a great effort in understanding the semantics and how it works together with substitution. This makes the work of proving the soundness theorem very simple. We will go ahead as we did for propositional logic.

13.1.1 Theorem (the soundness theorem)**.** *Consider a derivation in natural deduction. Then the conclusion is true in all interpretations where the undischarged assumptions are true.*

Proof. Remind yourself how the proof of the soundness theorem in propositional logic (6.1.5, page 45) went through. We will do a proof by induction according to exactly the same principles. We go through further cases now, depending on which rule is the last in the derivation $\mathcal{D}$. For the rules which already were present in Figure 5.1 (page 40), the treatment is exactly as in the proof of Theorem 6.1.5. We study the rules that were added in Figure 12.1 (page 90).

Case 10: $\mathcal{D}$ is of the form

$$\frac{}{t \doteq t}\,\text{refl} \tag{13.1.2}$$

We have $[\![t \doteq t]\!] = 1$, since $[\![t]\!] = [\![t]\!]$.

Case 11: $\mathcal{D}$ is of the form

$$\frac{\begin{array}{cc}\vdots & \vdots \\ \varphi[t/x] & t \doteq s\end{array}}{\varphi[s/x]\,.}\,\text{ers} \tag{13.1.3}$$

By the inductive hypothesis, it follows that $[\![\varphi[t/x]]\!] = 1$ and $[\![t \doteq s]\!] = 1$ in all interpretations where the undischarged assumptions are true. The first means, according to 11.2.25, that $[\![\varphi]\!]^{[x\mapsto[\![t]\!]]} = 1$, and the second means that $[\![t]\!] = [\![s]\!]$. Hence, we conclude that $[\![\varphi]\!]^{[x\mapsto[\![s]\!]]} = 1$, which gives $[\![\varphi[s/x]]\!] = 1$.

Note that the usage of Theorem 11.2.25 requires that t and s are free for x in φ.

Case 12: $\mathcal{D}$ is of the form

$$\frac{\begin{array}{c}\gamma_1 \ \cdots \ \gamma_n \\ \vdots \\ \varphi\end{array}}{\forall x\varphi}\,\forall I \tag{13.1.4}$$

where x does not occur freely in any undischarged assumption γ_i. We will show that in all interpretations with $[\![\gamma_i]\!] = 1$ for $i = 1, \ldots, n$, we have that $[\![\forall x\varphi]\!] = 1$, which means that $[\![\varphi]\!]^{[x\mapsto a]} = 1$ for every a in the domain. We can use the inductive hypothesis: it says that $[\![\varphi]\!]^{[x\mapsto a]} = 1$ holds if $[\![\gamma_i]\!]^{[x\mapsto a]} = 1$ holds for all $i = 1, \ldots, n$. But since x does not occur freely in γ_i, we have $[\![\gamma_i]\!]^{[x\mapsto a]} = [\![\gamma_i]\!] = 1$ (Theorem 10.1.28).

Case 13: $\mathcal{D}$ is of the form

$$\dfrac{\begin{array}{c}\vdots\\ \forall x\varphi\end{array}}{\varphi[t/x]}\;\forall E \tag{13.1.5}$$

Here we use that t is free for x.

By the inductive hypothesis, it follows that $\forall x\varphi$ is true in all interpretations in which the undischarged assumptions are true. Exercise 11.2.31 gives us, therefore, that $\varphi[t/x]$ is true as well in all such interpretations.

Case 14: $\mathcal{D}$ is fo the form

$$\dfrac{\begin{array}{c}\vdots\\ \varphi[t/x]\end{array}}{\exists x\varphi}\;\exists I \tag{13.1.6}$$

Here we use that t is free for x.

By the inductive hypothesis, it follows that $\varphi[t/x]$ is true in all interpretations in which the undischarged assumptions are true. Exercise 11.2.30 gives us, therefore, that $\exists x\varphi$ is true as well in all such interpretations.

Case 15: $\mathcal{D}$ is of the form

$$\dfrac{\begin{array}{c}\gamma_1 \cdots \gamma_n\\ \vdots\\ \exists x\varphi\end{array} \qquad \begin{array}{c}\gamma_{n+1} \cdots \gamma_m\ [\varphi]\\ \vdots\\ \sigma\end{array}}{\sigma}\;\exists E \tag{13.1.7}$$

where x does not occur freely in $\gamma_{n+1}, \ldots, \gamma_m$ nor in σ.

Note that we use both that x does not occur freely in $\gamma_{n+1}, \ldots, \gamma_m$ and that x does not occur freely in σ.

Take now an arbitrary interpretation $\mathcal{A}$ in which $\gamma_1, \ldots, \gamma_m$ is true. We will show that $[\![\sigma]\!]^{\mathcal{A}} = 1$. By the inductive hypothesis it follows that $[\![\exists x\varphi]\!] = 1$, which means that $[\![\varphi]\!]^{[x\mapsto a]} = 1$ for some a in the domain. Since x does not occur freely in $\gamma_{n+1}, \ldots, \gamma_m$, then, according to Theorem 10.1.28, we have $[\![\gamma_i]\!]^{[x\mapsto a]} = 1$ for $i = n+1, \ldots, m$. It follows by the inductive hypothesis, applied to the right subtree and to the interpretation $\mathcal{A}[x \mapsto a]$, that $[\![\sigma]\!]^{[x\mapsto a]} = 1$. But since x does not occur freely in σ, we have $[\![\sigma]\!] = [\![\sigma]\!]^{[x\mapsto a]} = 1$. □

13.1.8 Example. In the example 12.1.16 we first gave a wrong derivation of the formula $\exists x_0(x_0 \doteq x_1) \to \exists x_1(x_1 \doteq x_1)$ and a correct one afterwards, where a special rule for equality occurs. Now we are able to show that in fact one *must* use some of the special rules for equality, since these are the only rules which distinguish equality from other relations. In fact, it is imposible to derive $\exists x_0 P_1(x_0, x_1) \to \exists x_1 P_1(x_1, x_1)$, as we can realize through the help of the soundness theorem. Assume, indeed, that we had such a derivation without undischarged assumptions. This formula would also, according to the soundness theorem, be true in all interpretations. But if we proceed to interpret in the structure $\langle \mathbb{N}; >; \rangle$, then $\exists x_0 P_1(x_0, x_1)$ is interpreted as the proposition saying that there is a natural number greater than $v(x_1)$, which is true, while $\exists x_1 P_1(x_1, x_1)$ is interpreted as the proposition saying that there is a natural number greater than itself, which is false. Thus, the implication is false.

One can formulate the soundness theorem in the alternative way, also for predicate logic.

▶ **13.1.9 Definition** (cf. 6.1.16)**.** If $\Gamma \subseteq \text{Form}$, then $\Gamma \vDash \varphi$ means that every model of Γ is a model of φ.

▶ **13.1.10 Definition** (cf. 6.1.17)**.** If $\Gamma \subseteq \text{Form}$, then $\Gamma \vdash \varphi$ means that φ can be derived without any rules except those in figure 5.1 and 12.1, and without any undischarged assumptions except possible formulas in Γ.

The special case of the soundness theorem: $\vdash \varphi \Rightarrow \vDash \varphi$ says that only tautologies can be derived without undischarged assumptions.

13.1.11 Theorem (the soundness theorem in an alternative form)**.** $\Gamma \vdash \varphi \Rightarrow \Gamma \vDash \varphi$

Proof. Assume that $\Gamma \vdash \varphi$; that is to say, there is a derivation of φ where the undischarged assumptions $\gamma_1, \ldots, \gamma_n$ are all in Γ. All models of Γ are models of $\gamma_1, \ldots, \gamma_n$, and hence it follows from Theorem 13.1.1 that they are also models of φ, which was what we needed to show. □

The soundness theorem can also, among other things, be used to show that we cannot derive any new propositional formulas by use of the rules introduced in Figure 12.1.

Conservativity is a very important notion in mathematical foundations. The mathematician David Hilbert (1862–1943), who was so important that occupied more than two columns in the Swedish National Encyclopedia, thought that it should be the foundation of all justifications of advanced methods. What matters in the end, he said, is that specific theorems about simple computations were correct. To make it easier to reach such results, we could introduce "ideal elements" such as infinitely large numbers or other things to which mathematicians have gotten used. Reasoning about those ideal elements does not have to be "correct" in any other sense besides the fact that we should know that mathematics *with* such notions is conservative over mathematics *without* them. This became the foundation of what has been called *Hilbert's program*. Hilbert set as his goal to prove the conservativity of mathematics over the simpler "finitary" mathematics. Unfortunately mathematicians have not succeded. Today it is clear that we cannot complete Hilbert's program in the way Hilbert had in mind, and it is an open question whether the program can be modified in some reasonable way and be thereafter completed.

13.1.12 Theorem (conservativity)**.** *If $\Gamma \vdash \varphi$, and the formulas in Γ as well as φ are propositional, then there is a derivation that only uses the propositional rules (Figure 5.1).*

Proof. Assume that $\Gamma \vdash \varphi$. The soundness theorem gives us $\Gamma \vDash \varphi$. But it follows that $\Gamma \vDash \varphi$ holds even propositionally (since the interpretations of propositional formulas are the same as in propositional logic), and hence, according to the completeness theorem for propositional logic 8.2.3, we have that $\Gamma \vdash \varphi$ holds propositionally. □

Further definitions from propositional logic can be transferred directly to predicate logic

▶ **13.1.13 Definition.** By "Γ is inconsistent", we mean that $\Gamma \vdash \bot$. By "Γ is consistent", we mean that $\Gamma \nvdash \bot$.

13.1.14 Example. Show that $\{x_0 \doteq x_1, x_1 \doteq x_2, x_2 \doteq x_3, x_3 \doteq x_4\}$ is consistent.

Solution. Assume that $\{x_0 \doteq x_1, x_1 \doteq x_2, x_2 \doteq x_3, x_3 \doteq x_4\}$ is inconsistent; that is to say

$$x_0 \doteq x_1, x_1 \doteq x_2, x_2 \doteq x_3, x_3 \doteq x_4 \vdash \bot .$$

Then, according to the soundness theorem, we should have $x_0 \doteq x_1, x_1 \doteq x_2, x_2 \doteq x_3, x_3 \doteq x_4 \vDash \bot$. But with a valuation giving the same value to all variables we get a model of $x_0 \doteq x_1, x_1 \doteq x_2, x_2 \doteq x_3, x_3 \doteq x_4$ which is not a model of $\bot$ (no interpretation is), which shows that it is impossible for the set in question to be inconsistent. □

13.1.15 Exercise (from the exam on 2005-08-23, cf. Exercise 12.2.6)**.** Show that in fact it is impossible to derive $(\forall x\varphi \to \psi) \to (\exists x\varphi \to \psi)$, for certain choices of φ and ψ.

Hint. Consider the case $\psi = \bot$.

13.1.16 Exercise (from the exam on 2004-08-17, cf. Exercise 12.2.9)**.** Is there any correct way to derive $\forall x_0 \exists x_1 \varphi \to \exists x_1 \forall x_0 \varphi$ for all formulas φ?

13.1.17 Exercise. Show that $\varphi \vdash \forall x\varphi$ is not generally true.

13.1.18 Exercise (from the exam on 2004-10-18)**.** We have that $\varphi[y/x] \vdash \exists x\varphi$ with the help of a single instance of $\exists I$ if y is free for x in φ. Show that $\varphi \vdash \exists y\varphi[y/x]$ does not hold in general, but only if y is free for x in φ.

Hint. One can choose φ without quantifiers.

13.1.19 Exercise. Show that $\{\exists x_0 \neg(x_0 \doteq x_1)\} \cup \{x_0 \doteq x_1, x_1 \doteq x_2, x_2 \doteq x_3, \ldots\}$ is consistent.

13.1.20 Exercise (from the exam on 2002-08-20)**.** Let

$$\Gamma = \{\forall x_1 \exists x_2 P_1(x_1, x_2), \exists x_1 \forall x_2 P_1(x_1, x_2)\}$$

and $\varphi = \exists x_2 \forall x_1 P_1(x_1, x_2)$. Show that φ is independent of Γ, which means that $\Gamma \nvdash \varphi$ and $\Gamma \nvdash \neg\varphi$.

13.1.21 Exercise (from the exam on 2003-10-20)**.** Decide whether or not the following formula is derivable in natural deduction.

$$\forall x_0 \exists x_1 \neg(x_0 \doteq x_1)$$

13.1.22 Exercise (from the exam on 2003-10-20)**.** Show that we must use the assumption that x does not occur freely in ψ to be able to do Exercise 12.2.15.

Hint. Let, for instance, both φ and ψ be $x_0 \doteq x_1$.

Originally, Peano used nine axioms for the natural numbers, and they looked somewhat different. He considered 1 as the least natural number, but in the beginning of the 20th century it became more usual to include 0, and after some influential article in 1923, it became the dominant convention, at least within logic. Many of Peano's axioms are not needed in our presentation, since they can be derived through the rules of natural deduction.

13.1.23 Exercise (from the exam on 2004-10-18, cf. Exercise 12.2.11)**.** Is there, for every pair of formulas φ, ψ, a correct derivation of $\exists x_0(\varphi \wedge \psi) \to (\exists x_0 \varphi) \wedge \psi$?

13.1.24 Exercise (from the exam on 2004-01-08)**.** Peano's axioms for natural numbers are as follows. The language is assumed to contain a unary function symbol f_1 and a nullary function symbol f_2.

A1. $\neg\exists x_0(f_1(x_0) \doteq f_2)$

A2. $\forall x_0 \forall x_1(f_1(x_0) \doteq f_1(x_1) \to x_0 \doteq x_1)$

A3. $\varphi[f_2/x_0] \wedge \forall x_0(\varphi \to \varphi[f_1(x_0)/x_0]) \to \forall x_0 \varphi$

where A3 represents in fact infinitely many axioms, namely, one for every $\varphi \in$ Form.

Show that one cannot derive A1 from A2 and A3; that is, that there is no derivation of A1 where the undischarged assumptions are of the form A2 or A3.

13.2 Summary

We extended the proof of the soundness theorem to include the new rules, so that it is now proved for predicate logic. It turned out that the limitations on the variables which certain rules have is precisely what we need to apply the simplification rules for substitution in a way that helped to complete the proof of the soundness theorem. The most important thing to bring with you for the rest of the course is the ability to use the soundness theorem to detect when some ideas for constructing derivations are not fruitful, as well as showing whether a certain set of formulas is consistent.

13.3 Review exercises

13.3.1 Exercise (from the exam on 2003-01-09)**.** Let φ be the formula

$$\forall x_2(\forall x_1 P_1(x_1, x_2) \to \exists x_2(f_1(x_1) \doteq f_2(x_2, x_3))) \vee \forall x_3 \neg(x_1 \doteq x_3)\,.$$

a) Compute FV(φ)

b) Perform the substitutions $\varphi[f_1(x_3)/x_1]$, $\varphi[x_1/x_2]$, $\varphi[f_2(x_1, x_3)/x_3]$.

c) Specify, for each of the substitutions above, all of which are of the form $\varphi[t/x]$, whether t is free for x in φ.

13.3.2 Exercise (change of bound variables in $\forall$)**.**

a) Show that if y does not occur free in φ and y is free for x in φ, then:

$$\forall x \varphi \vdash \forall y \varphi[y/x]\,. \tag{13.3.3}$$

b) Give an example where y does not occur freely in φ but (13.3.3) does not hold.

c) Give an example where y is free for x in φ but (13.3.3) does not hold.

13.3.4 Exercise (change of bound variables in $\exists$)**.**

a) Show that if y does not occur freely in φ and y is free for x in φ, then:

$$\exists y \varphi[y/x] \vdash \exists x \varphi \,. \tag{13.3.5}$$

b) Give an example where y does not occur freely in φ but (13.3.5) does not hold.

c) Give an example where y is free for x in φ but (13.3.5) does not hold.

13.3.6 Exercise (from the exam on 2003-10-20)**.** In the Swedish National Encyclopedia[1] one can read the following under "Boolean algebra" (names and notation are changed to match those of this course):

> A Boolean algebra is defined as consisting of elements $a, b, c, \ldots$, which can be connected by Boolean operations $\vee$, $\wedge$ and $\neg$, so that $a \vee b$, $a \wedge b$ and $\neg a$ are elements of the algebra whenever a and b are. It is required that the following rules of computations (axioms) are fulfilled:
>
> 1) $a \vee b = b \vee a$ and $a \wedge b = b \wedge a$;
> 2) $a \wedge (b \vee c) = (a \wedge b) \vee (a \wedge c)$;
> 3) There are elements 0 and 1 such that $a \vee 0 = a \wedge 1 = a$ for all a;
> 4) $a \vee \neg a = 1$ and $a \wedge \neg a = 0$.
>
> $(\ldots)$ From these axioms one can derive more rules of computation, such as $(\ldots)$ the idempotence laws $a \vee a = a$ and $a \wedge a = a$.

Your task is to prove that the last claim is *wrong*. Do this in three steps:

a) Formalize axioms 1–4 in the language with arity type $\langle ; 2, 2, 1, 0, 0\rangle$, as formulas without free variables. Call them φ_1, φ_2, φ_3, φ_4.

b) Formalize the proposition: $a \vee a = a$ holds for all a. Call the resulting formula φ.

c) Show that $\varphi_1, \varphi_2, \varphi_3, \varphi_4 \not\vdash \varphi$.

Hint. Consider congruence modulo 2, i.e., define $+$ and $\cdot$ on $\{0, 1\}$ by the tables

$+$	0	1
0	0	1
1	1	0

$\cdot$	0	1
0	0	0
1	0	1

and define also an 1-ary operation ' by: $0' = 1$ and $1' = 0$.

13.3.7 Exercise (from the exam on 2004-08-17)**.** A structure $\langle A; ; \circ, e\rangle$ of arity type $\langle ; 2, 0\rangle$ is called a *monoid* if for all elements $a, b, c \in A$ we have:

$$\begin{aligned} a \circ e &= a \\ e \circ a &= a \\ a \circ (b \circ c) &= (a \circ b) \circ c \,. \end{aligned}$$

Examples of infinite monoids are $\langle \mathbb{N}; ; +, 0\rangle$ and $\langle \mathbb{N}; ; \cdot, 1\rangle$, where $\mathbb{N}$ are the natural numbers. Examples of finite monoids (with n elements) can be obtained from the previous one if we compute "modulo n": we consider numbers as equal if their difference is divisible by n.

[1] Nationalencyklopedin - A standard Swedish encyclopedia published during the period 1986-1996.

a) Formalize the definition; that is to say, give $\gamma_1, \gamma_2, \gamma_3 \in \text{Form}$ such that a structure is a monoid if and only if it is a model of $\gamma_1, \gamma_2, \gamma_3$.

b) Give a formula $\tau_1 \in \text{Form}$ which expresses that a monoid has only one element; that is to say, such that τ_1 is true in all such monoids, but false in all others.

c) Give a formula $\tau_2 \in \text{Form}$ which expresses that a monoid has exactly two elements.

13.3.8 Exercise (from the exam on 2004-01-08)**.** Decide, using the method of your preference, whether or not each of the following formulas is derivable through natural deduction. The language is assumed to contain 1-ary function symbols f_1 and f_2.

a) $\forall x_0 \forall x_1 \neg(x_0 \doteq x_1)$

b) $\forall x_0(\bot \to f_1(x_0) \doteq f_2(x_0))$.

13.3.9 Exercise (from the exam on 2004-08-17, cf. Exercises 12.2.9, 13.1.16)**.** Is there a correct way to derive $\exists x_1 \forall x_0 \varphi \to \forall x_0 \exists x_1 \varphi$ for all $\varphi \in \text{Form}$? Justify carefully!

13.3.10 Exercise (from the exam on 2004-01-08)**.** Interpret the formula:

$$\forall x_0 \exists x_1 (f_1(x_0, x_1) \doteq f_2)$$

in the following structure, and give its truth value in each one of them. Justify!

a) $\langle \mathbb{N};;+,0\rangle$

b) $\langle \mathbb{Z};;+,0\rangle$

c) $\langle \mathbb{R};;\cdot,1\rangle$

Here $\mathbb{N}$ are the natural numbers, $\mathbb{Z}$ the integers and $\mathbb{R}$ the real numbers.

13.3.11 Exercise. Show that if y is free for x in $\forall y \psi$ then $\psi[y/x] = \psi$.

13.3.12 Exercise (cf. Exercises 9.2.22, 11.2.37)**.** Show that $\varphi[y/x][x/y] = \varphi$ if y does not occur freely in φ and y is free for x in φ.

13.3.13 Exercise (similar to the exam problem from 2007-10-18)**.** Let $\psi = \forall x_0 \varphi \to \forall x_1 \varphi[x_1/x_0]$.

a) Show that if x_1 is free for x_0 in φ and x_1 does not occur freely in φ, then ψ is a tautology.

b) Give an example of φ in which x_1 is free for x_0 such that ψ is not a tautology.

c) Give an example of φ in which x_1 does not occur freely such that ψ is not a tautology.

d) Give an example of φ in which x_1 is bound for x_0 and x_1 occurs freely such that ψ is a tautology.

13.3.14 Exercise.

a) What is wrong in the following derivation?

$$\dfrac{[\exists x_0(x_0 \doteq x_1)]^1 \qquad \dfrac{[x_0 \doteq x_1]^2}{\exists x_1(x_1 \doteq x_1)}\,\exists I}{\dfrac{\exists x_1(x_1 \doteq x_1)}{\exists x_0(x_0 \doteq x_1) \to \exists x_1(x_1 \doteq x_1)}\,{\to}I_1}\,\exists E_2$$

b) Can we derive $\exists x_0(x_0 \doteq x_1) \to \exists x_1(x_1 \doteq x_1)$?

c) Can we derive $\exists x_0 P_1(x_0, x_1) \to \exists x_1 P_1(x_1, x_1)$?

Chapter 14

Completeness

We will now prove the completeness theorem for predicate logic. The setup for this will be like that for propositional logic, but with more intricate details. It is often said that this was first proved by Gödel in 1930 in his PhD thesis, but the Norwegian mathematician Skolem already proved it in 1922.

14.1 Maximal consistency and the existence property

Several definitions and theorems about maximal consistency for propositional logic can be transferred directly to predicate logic.

▶ **14.1.1 Definition** (cf. Definition 8.1.1)**.** Γ is *maximally consistent* provided it is maximal amongst consistent subsets of Form with respect to inclusion. In other words, it means that:

1. Γ is consistent,
2. If $\Gamma \subseteq U \subseteq \text{Form}$ and U is consistent, then $U = \Gamma$.

14.1.2 Theorem. *If* Γ *is maximally consistent and* $\Gamma \vdash \varphi$*, then* $\varphi \in \Gamma$*.*

Proof. Same proof as Theorem 8.1.2. ☐

It is good if you make sure to practice these results, e.g., by trying to prove it by yourself.

14.1.3 Theorem. Γ *is maximally consistent if and only if it is consistent and whenever* $\Gamma \cup \{\varphi\}$ *is consistent, then* $\varphi \in \Gamma$*.*

Proof. Same proof as Theorem 8.1.4. ☐

14.1.4 Exercise (from the exam on 2005-01-07)**.** Assume that Γ is a maximally consistent set of formulas.

a) Give an example of a formula that has to be in Γ. Motivate!

b) Show that if $\exists x_1 P_1(x_1) \in \Gamma$, then $\exists x_2 P_1(x_2) \in \Gamma$.

14.1.5 Exercise (from the exam on 2003-10-20)**.** Let $\Gamma = \{P_1(x_0), P_1(x_1), P_1(x_2), \ldots\}$.

a) Show that Γ is consistent but not maximally consistent.

b) Is Γ complete? I.e., is it true that for every formula φ without free variables $\Gamma \vdash \varphi$ or $\Gamma \vdash \neg\varphi$?

c) Let Γ^* be maximally consistent and $\Gamma \subseteq \Gamma^*$. Show that $\exists x_0 P_1(x_0) \in \Gamma^*$.

14.1.6 Exercise (cf. Theorem 8.1.11)**.** If Γ is maximally consistent and $\varphi \notin \Gamma$, then $\neg\varphi \in \Gamma$.

For propositional logic, we showed how to extend consistent sets to maximally consistent sets. This was used later to prove that consistent sets have models. Also in predicate logic such an extension can be made with exactly the same method, though it is not enough to have that for the proof of the model existence lemma; you also need the extension to satisfy the *existence property.* This means that if $\exists x\psi$ is in the set, then $\psi[t/x]$ is also in the set for some term t free for x in ψ. Therefore, we need to modify the construction of Γ^* somewhat. We will need infinitely many variables not occurring freely in Γ. Those will always exists if Γ is finite, but if we had $\{x_0 \doteq x_1,\ x_1 \doteq x_2,\ x_2 \doteq x_3,\ \ldots\}$ then all variables would occur freely. We first treat the case where there are enough variables to work with; the other case will be handled later.

14.1.7 Lemma (Maximal consistent extension with the existence property)**.** *Let Γ be consistent and suppose there are infinitely many variables that do not occur freely in Γ. Then there is a maximal consistent extension Γ^* which has the following existence property: if a formula of the form $\exists x\psi$ belongs to the extension, then also $\psi[t/x]$ belongs to the extension, for some term t free for x in ψ.*

You call t a *witness* for the existential formula

Proof. Let $\{\varphi_0, \varphi_1, \varphi_2, \ldots\}$ be an enumeration of Form. We will, as in propositional logic, go through this list and for each formula we will decide whether it will belong to Γ^*. We therefore construct a growing sequence $\{\Gamma_n\}$ of subsets of Form, where $\Gamma_0 = \Gamma$ and the union of all of them is Γ^*. We define it as follows:

$$\Gamma_0 \stackrel{\text{def}}{=} \Gamma$$

$$\Gamma_{s(n)} \stackrel{\text{def}}{=} \begin{cases} \Gamma_n \cup \Gamma_n' & \text{if } \Gamma_n \cup \{\varphi_n\} \text{ consistent} \\ \Gamma_n & \text{otherwise.} \end{cases}$$

Here, $\Gamma_n' = \{\varphi_n\}$ always except when φ_n is of the form $\exists x\psi$. In that case we let $\Gamma_n' = \{\exists x\psi,\ \psi[y/x]\}$, where y is a variable chosen in a way that it does not occur freely in any formula in Γ_n and does not occur at all in ψ.

To be able to chose such a y is the reason why we asked for infinitely many variables.

Now let

$$\Gamma^* \stackrel{\text{def}}{=} \bigcup_{n=0}^{\infty} \Gamma_n\,. \tag{14.1.8}$$

We shall check that Γ^* has the required property. To check that Γ^* is consistent, as in predicate logic, we just make sure that every Γ_n is consistent (look at the proof of Theorem 8.1.10), which is done inductively. That Γ_0 is consistent follows from the fact that $\Gamma_0 = \Gamma$. As the induction step, we will show that $\Gamma_{s(n)}$ is consistent if Γ_n is consistent. If φ_n is not of the form $\exists x\psi$, it is obvious that $\Gamma_{s(n)}$ is consistent, since $\Gamma_{s(n)}$ is chosen to be a consistent set. We must handle the case where φ_n is of the form $\exists x\psi$. We shall prove that if $\Gamma_n \cup \{\exists x\psi\}$ is consistent, then $\Gamma_n \cup \Gamma_n'$ is also consistent. Assume therefore that $\Gamma_n \cup \{\exists x\psi\}$ was consistent but that we had a derivation of $\bot$ from $\Gamma_n \cup \{\exists x\psi,\ \psi[y/x]\}$. Then the derivation could be modified in the following way:

$$\dfrac{\dfrac{\exists x\psi \qquad \dfrac{[\psi]}{\exists y\psi[y/x]}\,\exists I}{\exists y\psi[y/x]}\,\exists E \qquad \begin{matrix}\Gamma_n \quad \exists x\psi \quad [\psi[y/x]] \\ \vdots \\ \bot\end{matrix}}{\bot}\,\exists E \tag{14.1.9}$$

and we would therefore also have a derivation of $\bot$ from $\Gamma_n \cup \{\exists x\psi\}$, which is impossible by assumption. We must however check that the derivation is correct. The application of the $\exists I$ is correct since $\psi[y/x][x/y] = \psi$, because y does not occur in ψ (Exercise 9.2.22) and x is free for y in $\psi[y/x]$ (Exercise 11.2.38). The application of $\exists E$ in the row underneath is correct, since x does not occur freely in $\exists y\psi[y/x]$. Finally, the last application of $\exists E$ is correct, since y was chosen so that it does not occur freely in any formula in Γ_n, nor in ψ.

We know that Γ^* is consistent, but we need to know that it is *maximally* consistent and that satisfies the existence property. But it follows from Theorem 14.1.3 that if $\Gamma^* \cup \{\varphi_n\}$ is consistent, so is $\Gamma_n \cup \{\varphi_n\}$, and hence $\Gamma'_n \subseteq \Gamma^*$ (since $\Gamma'_n \subseteq \Gamma_n \cup \Gamma'_n = \Gamma_{s(n)} \subseteq \Gamma^*$). This gives us both maximal consistency and the existence property, because Γ'_n has been constructed to meet two needs: it always contains φ_n when $\Gamma^* \cup \{\varphi_n\}$ is consistent, but also $\psi[y/x]$ when φ_n is of the form $\exists x\psi$. □

14.1.10 Exercise. Let Γ consist of the formulas (in the language of arity type $\langle ; 2, 0\rangle$) which are true in the structure $\langle \mathbb{Z}; ; +, 0\rangle$ if we use the interpretation $v(x_i) \stackrel{\text{def}}{=} i$.

a) Give an example of a formula of Γ containing two different variables but no quantifiers.

b) Show that Γ is maximally consistent.

c) Does Γ have the existence property?

Hint. Use that all terms are non-negative values in the current interpretation.

14.2 Completeness

We will construct a model $\mathcal{A}$ of Γ^*. The idea is to interpret the language as referring to their own terms. We shall thus interpret $\exists$-formulas as saying that there is a term with a particular property, and so on. The formula $t \doteq s$ shall therefore say that the terms t and s are alike. This does not really work, because we should not consider each Term individually, but divide the set of terms in equivalence classes given by the equivalence relation

$$t \sim s \stackrel{\text{def}}{=} \big((t \doteq s) \in \Gamma^*\big)\,. \tag{14.2.1}$$

14.2.2 Exercise. Show that $\sim$ is an equivalence relation.

Let $|\mathcal{A}|$ be the set of equivalence classes. We will try to have each term interpreted by its own equivalence class, and we will note $v(x_i)$ for the equivalence class containing x_i. We will interpret, furthermore, each function symbol f_i, by the function $f_i^{\mathcal{A}}$ from the equivalence classes of $t_1, \dots, t_{a_i}$ to the equivalence classes of the terms $f_i(t_1, \dots, t_{a_i})$. If for a term t we denote its equivalence class as $\tilde{t}$, we can define the interpretation as follows:

$$v(x_i) = \tilde{x_i}$$
$$f_i^{\mathcal{A}}(\tilde{t_1}, \dots, \tilde{t_{a_i}}) = \widetilde{f_i(t_1, \dots, t_{a_i})}$$

14.2.3 Exercise. Show that the functions $f_i^{\mathcal{A}}$ are *well defined*; that is, that their values do not depend on the choice of representatives of each equivalence class: if $t_j \sim s_j$ for $j = 1, \dots, a_i$, then $f_i(t_1, \dots, t_{a_i}) \sim f_i(s_1, \dots, s_{a_i})$.

Hint. Use that Γ^* is closed under derivations, and that the replacement rule can be used to derive $f_i(t_1, \dots, t_{a_i}) \doteq f_i(s_1, \dots, s_{a_i})$ from the formulas $t_j \doteq s_j$.

14.2.4 Exercise. Show that $[\![t]\!] = \tilde{t}$ for each term t.

For this reason, we will not use the notation $\tilde{t}$ in what follows, but will instead use $[\![t]\!]$.

Hint. We know that $[\![t]\!]$ is an equivalence class, since individuals are interpreted as equivalence classes. What we need to show is that it is the "right" equivalence class. If t is the variable x_i, this follows easily, since $[\![x_i]\!] = v(x_i)$, but what is $[\![t]\!]$ when t is not a variable? To show that this holds for all cases use, induction in the construction of the terms.

14.2.5 Exercise. Show that $t \in [\![t]\!]$ for each term t.

The next lemma shows that each individual can be represented in our interpretation by a particularly useful term.

14.2.6 Lemma. *For each individual a and each choice of φ and x, there exists a term t such that $a = [\![t]\!]$ and t is free for x in φ.*

See Exercise 14.2.17 for an alternate proof.

Proof. Take $s \in a$. Then $a = [\![s]\!]$, according to the previous exercise. There are an infinite number of variables z that do not occur in s, and for each of these, the formula $\exists z(z \doteq s)$ is derivable, so Γ^* contains infinitely many such formulas. For each one of them, there exists a variable y, chosen in the construction of Γ^*, such that $(y \doteq s) \in \Gamma^*$. Since these variables are all different, there must be one amongst them that is free for x in φ. Take t as such a variable. □

We now define the interpretation of the formulas as follows:

$$P_i^{\mathcal{A}}([\![t_1]\!], \dots, [\![t_{r_i}]\!]) \stackrel{\text{def}}{=} \big(P_i(t_1, \dots, t_{r_i}) \in \Gamma^*\big)\,. \tag{14.2.7}$$

We need to check that these interpretations are well defined in a similar sense as in Exercise 14.2.3: that they do not depend on the choice of the representative of each equivalence class. We will skip the details since those are similar to the ones in the mentioned exercise.

We will now verify that we have really constructed a model.

Lemma 14.2.8 and the exercise that follows show that we have managed to interpreted formulas that only involve terms. An $\forall$-formula is interpreted as true precisely when all terms satisfy a corresponding property, while an $\exists$-formula is interpreted as true precisely when at least one term satisfies the property. Funnily enough, since these results do not rely on how we chose to interpret relation symbols!

14.2.8 Lemma. $[\![\forall x\varphi]\!] = 1$ *is equivalent to having* $[\![\varphi[t/x]]\!] = 1$ *for all terms t that are free for x in φ.*

Proof. $\forall x\varphi \vDash \varphi[t/x]$ has been shown in Exercise 11.2.31. We shall prove the other implication. Assume, therefore, that $[\![\varphi[t/x]]\!] = 1$ for all terms t that are free for x in φ. For such terms, we also have $[\![\varphi]\!]^{[x\mapsto[\![t]\!]]} = 1$. But according to the previous lemma, each individual is of the form $[\![t]\!]$ for that kind of terms, so we have $[\![\forall x\varphi]\!] = 1$. □

14.2.9 Exercise. Show that $[\![\exists x\varphi]\!] = 1$ is equivalent to $[\![\varphi[t/x]]\!] = 1$ for some term t that is free for x in φ.

Hint. Use the previous lemma.

14.2.10 Lemma. *For any formula φ we have:* $[\![\varphi]\!] = 1 \iff \varphi \in \Gamma^*$.

Proof. We will prove this by induction, though now the induction will be done in the *number of logical operations* in the formula (we defined the number of logical operations in Exercise 6.3.3, but now we will also take into account $\forall$ and $\exists$ as logical operations). We therefore consider the following statement:

Note the informal restricted quantifier!

> For each natural number n and all formulas φ containing n logical operations, we have $[\![\varphi]\!] = 1 \iff \varphi \in \Gamma^*$.

To carry on the proof, we simply have to go through the various forms a formula can have, and use, in each step, the inductive hypothesis, which is:

> For each formula φ' with fewer logical operations than φ we have $[\![\varphi']\!] = 1 \iff \varphi' \in \Gamma^*$.

In the case of equalities of terms it is easy to prove the statement, and we do not need to consider the inductive hypothesis:

$$[\![t_1 \doteq t_2]\!] = 1 \iff [\![t_1]\!] = [\![t_2]\!] \iff t_1 \sim t_2 \iff (t_1 \doteq t_2) \in \Gamma^*. \tag{14.2.11}$$

In the case of relation symbols, the proof is just as easy:

$$[\![P_i(t_1, \dots, t_{r_i})]\!] = 1 \iff P_i^{\mathcal{A}}([\![t_1]\!], \dots, [\![t_{r_i}]\!]) \iff P_i(t_1, \dots, t_{r_i}) \in \Gamma^*. \tag{14.2.12}$$

The connectives are handled in the same way as in predicate logic (proof of Lemma 8.2.2). For formulas of the form $\forall x\psi$, note that, according to the previous lemma, $[\![\forall x\psi]\!] = 1$ is equivalent to $[\![\psi[t/x]]\!] = 1$ for all terms t which are free for x in ψ. By the inductive hypothesis, this happens to be equivalent to $\psi[t/x] \in \Gamma^*$ for all terms t that are free for x in ψ. This is, in turn, equivalent to having $\forall x\psi \in \Gamma^*$. Indeed, to see one implication we note that $\forall x\psi \vdash \psi[t/x]$ and that Γ^* is closed under derivations. To check the other implication, we reason as follows: we cannot have $\exists x\neg\psi \in \Gamma^*$, since then the existence property would give us $\neg\psi[t/x] \in \Gamma^*$ for a term t which is free for x in ψ, and this is not possible since then $\psi[t/x] \in \Gamma^*$ for such terms, while Γ^* is consistent. Therefore, it follows from Exercise 14.1.6 that $\neg\exists x\neg\psi \in \Gamma^*$. Since Γ^* is closed under derivations, we have $\forall x\psi \in \Gamma^*$.

Here we use the existence property.

Finally, let us consider the case of the formula $\exists x\psi$. According to the previous exercise, $[\![\exists x\psi]\!] = 1$ is equivalent to having $[\![\psi[t/x]]\!] = 1$ for any term t that is free for x in ψ. But $\psi[t/x]$ has one less logical operation than $\exists x\psi$, so we can apply the inductive hypothesis and conclude that $[\![\psi[t/x]]\!] = 1$ is equivalent to $\psi[t/x] \in \Gamma^*$. This, in turn, is easily seen to be equivalent to $\exists x\psi \in \Gamma^*$. Indeed, one implication follows from the fact that Γ^* is closed under derivations and $\psi[t/x] \vdash \exists x\psi$, while the other direction follows from the existence property. □

14.2.13 Lemma (Model existence)**.** *Every consistent subset of* Form *has a model.*

Proof. Suppose that Γ is a consistent set of Form. According to the previous lemma, it is enough to extend that set to Γ^*, since then we can find a model for Γ^*, which will also be a model for Γ. However, if Γ is infinite and has infinitely many free variables, the construction of Γ^* cannot always be performed, and we must solve this issue.

Construct a different set Γ' by replacing in each formula of Γ the variable x_i by x_{2i}. Since no variables with odd index occur in Γ', the construction of ${\Gamma'}^*$ works, and hence Γ' has a model. The same interpretation is also a model of Γ provided we change the valuations to match the change of variables we performed. More specifically, if v' is the valuation corresponding to the constructed model of Γ', we can therefore put $v(x_i) = v'(x_{2i})$ and obtain a model of Γ. □

14.2.14 Theorem (completeness theorem)**.** *If* $\Gamma \vDash \varphi$*, then* $\Gamma \vdash \varphi$*.*

Proof. Is similar to the proof of 8.2.3. □

14.2.15 Exercise. We can derive the completeness theorem relatively easily from the model existence lemma. Conversely, it is possible to derive that lemma from the completeness theorem. Find out how.

14.2.16 Exercise (from the exam on 2003-10-20)**.** Determine whether the following formula is derivable through natural deduction:

$$\forall x_0 \exists x_1 \forall x_2 (x_0 \doteq x_1 \to x_1 \doteq x_2)$$

14.2.17 Exercise (optional)**.** The given proof of lemma 14.2.6 uses the construction of Γ^*. Changing the construction would therefore force us to modify this proof. It is then of interest to find a proof that only relies on the assumptions that Γ^* is maximally consistent and has the existence property. Here we outline such a proof. The exercise consists of carrying it out in detail.

Take $s \in a$ and let $y_1, \dots, y_n$ the variables bound by any quantifier in φ. We will show that there is a term t such that $(t \doteq s) \in \Gamma^*$ and the variables $y_1, \dots, y_n$ do not appear in t (why is this enough?). Take, therefore, a variable y different from $y_1, \dots, y_n$ and that does not appear in s. The formula:

$$\exists y(y \doteq s \wedge \forall y_1 \cdots \forall y_n (y \doteq y)) \qquad (14.2.18)$$

is then derivable (how?) and hence it is in Γ^* (why?). By the existence property, it follows that there is a term t that is free for y in $(y \doteq s \wedge \forall y_1 \cdots \forall y_n(y \doteq y))$ and such that

$$(y \doteq s \wedge \forall y_1 \cdots \forall y_n(y \doteq y))[t/y] \in \Gamma^* . \tag{14.2.19}$$

It follows that $y_1, \ldots, y_n$ do not appear in t (why?) and that $(t \doteq s) \in \Gamma^*$ (why?).

The compactness theorem has a topological meaning. Consider two interpretations as similar when the same set of formulas is true in both of them. This divides the set of interpretations in equivalence classes, which can be considered as points of a topological space. Let each formula represent the set of interpretations (up to equivalence) that satisfy it. Then Form is a base of closed sets of a topology. The theorem then says that a family of closed sets has nonempty intersection provided each finite subfamily has nonempty intersection. Thus, the space is compact in a topological sense.

14.3 Compactness

Another application of the existence lemma is the following remarkable theorem.

14.3.1 Theorem (Compactness theorem)**.** Γ *has a model if and only if every finite subset of* Γ *has a model.*

Proof. ($\Rightarrow$) If $\mathcal{A}$ is a model of Γ, then $\mathcal{A}$ is also a model of every finite subset of Γ.

($\Leftarrow$) Suppose that every finite subset of Γ has a model. Then every finite subset is consistent, by the soundness theorem. Hence, Γ is itself consistent. By the model existence lemma it follows now that Γ has a model. □

14.3.2 Example (non-standard numbers)**.** We can show that there is a model of Peano's axioms in which there are "infinite" numbers.

Peano's axioms for natural numbers are the following:

A1. $\neg\exists x_0(f_1(x_0) \doteq f_2)$

A2. $\forall x_0 \forall x_1(f_1(x_0) \doteq f_1(x_1) \to x_0 \doteq x_1)$

A3. $\varphi[f_2/x_0] \wedge \forall x_0(\varphi \to \varphi[f_1(x_0)/x_0]) \to \forall x_0 \varphi$

where A3 actually represents an infinite number of formulas, one for each $\varphi \in$ Form.

Let Γ consist of Peano's axioms together with the following formulas:

$$\begin{aligned}
\varphi_0 &= \neg(x_0 \doteq f_2) \\
\varphi_1 &= \neg(x_0 \doteq f_1(f_2)) \\
\varphi_2 &= \neg(x_0 \doteq f_1(f_1(f_2))) \\
\varphi_3 &= \neg(x_0 \doteq f_1(f_1(f_1(f_2)))) \\
&\vdots
\end{aligned}$$

(infinitely many). It is clear that the natural numbers are not a model of Γ (if f_1 is interpreted as s and f_2 as 0) since x_0 cannot be valued in a way that all formulas are true: the formula $\varphi_{v(x_0)}$ will be false. However, there exists another model. It can be shown as follows:

According to the compactness theorem, it suffices to show that every finite subset of Γ has a model. Take therefore a finite subset Γ_0 of Γ. Choose $n \in \mathbb{N}$ as the largest number such that $\varphi_n \in \Gamma_0$. Interpret this theory in the natural numbers and let $v(x_0) > n$. With this interpretation, we have $[\![\varphi_i]\!] = 1$ for every $i \leq n$, and hence it is a model of Γ_0. But Γ_0 was an arbitrary finite subset of Γ, so every finite subset of Γ has a model. Therefore, Γ has a model.

14.3.3 Exercise (from the exam on 2004-01-08)**.** Let Γ consists of Peano's axioms together with the formulas:

$$
\begin{array}{l}
P_1(f_2)\\
P_1(f_1(f_2))\\
P_1(f_1(f_1(f_2)))\\
\quad\vdots\\
P_1(\underbrace{f_1(f_1(\cdots}_{n \text{ stycken } f_1} f_2\cdots)))\\
\quad\vdots
\end{array}
$$

(one formula for each natural number n).

Does $\Gamma \cup \{\exists x_0 \neg P_1(x_0)\}$ have a model?

14.3.4 Exercise (from the exam on 2004-08-17).

a) Suppose that $\gamma_1, \gamma_2, \gamma_3$ solve exercise 13.3.7 a. Assume also that $\varphi \in$ Form is true in all monoids. Is it safe to say that $\gamma_1, \gamma_2, \gamma_3 \vdash \varphi$?

b) Recall Exercises 13.3.7 b and 13.3.7 c. Is there any formula $\tau \in$ Form expressing that a monoid is *finite* (that is, such that τ is true in all finite monoids but false in all infinite ones)? Explain carefully!

There is a part of logic called *model theory*. It studies the properties of models of different theories, as well as the theories whose models have specific characteristics. One typical question is to find what type of theories have finite models, countable models, etc.

14.4 Summary

We have gone through the concept of *maximal consistency* in predicate logic and proved that every consistent set can be extended to a maximally consistent set. This was still not good enough to construct a model of a consistent set, so we have also introduced the concept of the *existence property*. We saw that all consistent sets may be extended to maximally consistent sets that have the existence property, and showed how this, in turn, can be used to prove that *all* consistent sets have models (not just those that can be extended as explained). This allowed us to prove the completeness theorem for predicate logic. The theorem shows that the system contains all the rules necessary to derive valid formulas. If a formula cannot be derived in our system, then it is false in some interpretation. Finally we also studied the *compactness*. Using this concept we were able to construct models for infinite sets of formulas by looking at the models for finite subsets of them, which is usually considerably easier.

It is important that you understand what the completeness theorem says, and how it can be used to show that some formula can be derived without actually constructing the explicit derivation.

We hope you have enjoyed the course!

Part IV

Appendix and index

Normalization proofs

Proof of Glivenko's theorem (7.2.3)

The proof is taken from Seldin[a], with small adjustments to fit our system.

[a] Seldin, J. Normalization and Excluded Middle I, in Studia Logica 48, pp. 193-217, 1989.

We shall proof that if we use RAA further up in a derivation than the last step, one can change the derivation so that the usage of RAA is pushed down. By doing this repeatedly one will get in the end a derivation where RAA is not used except, possible, at the last step.

Assume therefore that RAA is used a little further up in the derivation. Call the following rule R, so that the derivation has the following form:

$$\dfrac{\dfrac{\begin{matrix}[\neg\varphi]^1\\ \vdots\\ \bot\end{matrix}}{\varphi}\,\text{RAA}_1 \qquad \vdots}{\psi}\,R \tag{A.1}$$

The vertical dots to the right, next to R, denote other possible subderivations which exist above R. If R is a rule with only one premise, no such subderivations exists, so the dots can be taken out, but if R has more premises (one or two more), the derivations of these will be placed where the dots are. If now R is any rule which does not discharge an assumption in the derivation of φ, we transform the derivation in the following way:

$$\dfrac{\begin{matrix}\dfrac{\dfrac{[\neg\psi]^1 \qquad \dfrac{[\varphi]^2 \qquad \vdots}{\psi}\,R}{\bot}\,{\to}E}{\neg\varphi}\,{\to}I_2\\ \vdots\\ \bot\end{matrix}}{\psi}\,\text{RAA}_1 \tag{A.2}$$

Note that the usage of RAA is pushed downwards. If on the other hand R discharges an assumption in the derivation of φ, then we cannot transform in this way, since R can no longer do the discharge. We must therefore handle these cases one by one. Only three rules discharge assumptions:

Case $\vee E$:

$$\dfrac{\begin{matrix}\vdots\\ \varphi\vee\psi\end{matrix} \qquad \dfrac{\begin{matrix}[\neg\sigma]^1 \quad [\varphi]^2\\ \vdots\\ \bot\end{matrix}}{\sigma}\,\text{RAA}_1 \qquad \begin{matrix}[\psi]^2\\ \vdots\\ \sigma\end{matrix}}{\sigma}\,{\vee}E_2 \tag{A.3}$$

A derivation of this kind can be transformed in the following way: one replaces

RAA with $\bot E$ and then concludes by using RAA:

$$\dfrac{\dfrac{[\neg\sigma]^1 \qquad \dfrac{\begin{matrix}\vdots\\ \varphi\vee\psi\end{matrix} \qquad \dfrac{\begin{matrix}[\neg\sigma]^1\quad[\varphi]^2\\ \vdots\\ \bot\end{matrix}}{\sigma}\,{\scriptstyle \bot E} \qquad \begin{matrix}[\psi]^2\\ \vdots\\ \sigma\end{matrix}}{\sigma}\,{\scriptstyle \vee E_2}}{\bot}\,{\scriptstyle \to E}}{\sigma}\,{\scriptstyle \text{RAA}_1} \tag{A.4}$$

One proceeds similarly if RAA occurs as the last rule in the right subderivation or in both subderivations.

Case $\to I$:

$$\dfrac{\dfrac{\begin{matrix}[\neg\psi]^1\qquad[\varphi]^2\\ \vdots\\ \bot\end{matrix}}{\psi}\,{\scriptstyle \text{RAA}_1}}{\varphi\to\psi}\,{\scriptstyle \to I_2} \tag{A.5}$$

We transform this to:

$$\dfrac{\dfrac{[\neg(\varphi\to\psi)]^1 \qquad \dfrac{\dfrac{\begin{matrix}\dfrac{\dfrac{[\neg(\varphi\to\psi)]^1 \qquad \dfrac{[\psi]^3}{\varphi\to\psi}\,{\scriptstyle \to I}}{\bot}\,{\scriptstyle \to E}}{\neg\psi}\,{\scriptstyle \to I_3} \qquad [\varphi]^2\\ \vdots\\ \bot\end{matrix}}{\psi}\,{\scriptstyle \bot E}}{\varphi\to\psi}\,{\scriptstyle \to I_2}}{\bot}\,{\scriptstyle \to E}}{\varphi\to\psi}\,{\scriptstyle \text{RAA}_1} \tag{A.6}$$

Case RAA: This is quite strange. No one that is somewhat experienced will derive in the following way, but for the sake of completeness we must cover also this case. Assume, then, that we have a derivation of the following form:

$$\dfrac{\dfrac{\begin{matrix}[\neg\sigma]^2\qquad[\neg\bot]^1\\ \vdots\\ \bot\end{matrix}}{\bot}\,{\scriptstyle \text{RAA}_1}}{\sigma}\,{\scriptstyle \text{RAA}_2} \tag{A.7}$$

Even if such a derivation is not constructed manually, this sort of derivation can in fact occur when one uses the transformations we have gone through above. In such situations, we transform by replacing the assumption of $\neg\bot$ with derivations of such formulas, so that the first RAA step can be completely removed. The same technique can be used in any case where the conclusion in RAA is $\bot$:

$$\dfrac{\begin{matrix}[\neg\bot]\\ \vdots\\ \bot\end{matrix}}{\bot}\,{\scriptstyle \text{RAA}} \tag{A.8}$$

transforms into

$$\begin{matrix}\dfrac{[\bot]}{\neg\bot}\,{\scriptstyle \to I}\\ \vdots\\ \bot\end{matrix} \tag{A.9}$$

By using the transfromations we have mentioned above, one can move the usage of RAA further and further down in the derivation, so that in the end there is at most one usage: as the bottom most rule. We must however check that this process really comes to an end. It might very well happen that we get *more* applications of RAA when we make transformations from (A.1) to (A.2), namely, if the dotted subderivation by the rule R is copied several times and contains RAA.

We therefore do a proof by induction over the structure of derivations. The inductive hypothesis is hence that the the theorem is true for all subderivations in the last rule, and we shall now prove the theorem for the whole derivation.We then have to consider the cases for which we formulated the transformation principles above. All these cases are simple to handle, except the first one: from (A.1) to (A.2). We go over this case. The inductive hypothesis is then that the subderivations in (A.1) above the rule R do not contain RAA except possibly as the last rule. After the transformation we know, therefore, that the derivation looks like (A.2), and in addition to the shown occurrence of RAA, it can only occur as the last rule in the uppermost dotted part. Consider now the subderivation which contains this dotted part and extends a couple of steps further down, with $\bot$ as its conclusion. This subderivation contains at most one occurrence of RAA, and can therefore be transformed in the way stipulated by the theorem. But then one can get rid of RAA from this part, since RAA, whose conclusion is $\bot$, can be removed by one of the transformation principles. The conclusion is that, one after one, the usages of RAA are removed, until only the bottom most is left. □

Proof of weak normalization (7.2.6)

We will show how, by a number of transformation rules, we can transform a derivation into a normal derivation. We first present the various transformations. Later, we will check that we can do the transformation process in such a way that we are certain that we will eventually reach a normal derivation. While you read the transformation principles, you should note that no derivation rules are added. Sometimes subderivations are copied several times (namely, when more assumptions of the same formula are replaced by subderivations), so the number of usages of a certain derivation rule can increase, but a derivation rule which is not used in the original derivation cannot occur in the resulting derivation either.

If we have the form

$$\dfrac{\dfrac{\begin{matrix}\vdots & \qquad & \vdots\\ \varphi & \qquad & \psi\end{matrix}}{\varphi\wedge\psi}\,{\scriptstyle\wedge I}}{\varphi}\,{\scriptstyle\wedge E} \tag{A.10}$$

we reduce to the left subderivation:

$$\begin{matrix}\vdots\\ \varphi\end{matrix} \tag{A.11}$$

and by using the other and-elimination rule we reduce to the right subderivation.

Linguistic expressions such as "taking out detours" can give the impression that the result is "better" in some sense. It is true that awkwardly constructed derivations can often be simplified through normalization, but it is also often the case that normalization increases the length of derivations. It is in this sense that they can become "worse". In the beginning, it can however be good to think of normalization as simplification.

If we have the form:

$$\dfrac{\dfrac{\begin{matrix}\vdots\\ \varphi\end{matrix}}{\varphi\vee\psi}\,{\scriptstyle\vee I} \qquad \begin{matrix}[\varphi]\\ \vdots\\ \sigma\end{matrix} \qquad \begin{matrix}[\psi]\\ \vdots\\ \sigma\end{matrix}}{\sigma}\,{\scriptstyle\vee E} \tag{A.12}$$

we reduce to:

$$\begin{array}{c}\vdots\\ \varphi\\ \vdots\\ \sigma\end{array} \tag{A.13}$$

and similarly for the other or-introduction rule. In the figure, it looks as if we were simplifying. In fact, the derivation can grow explosively by such reduction. The reason for this is that *every* occurrence of the discharged assumption $[\varphi]$ is replaced by a derivation of φ. Since there may be many such occurrences and the derivation we insert may be very long, we can get very big derivations as the result of the reduction. What we gain is that the derivation will become one step closer to being normal.

If we have the form:

$$\dfrac{\dfrac{\begin{array}{c}[\varphi]\\ \vdots\\ \psi\end{array}}{\varphi \to \psi}\,{\to}I \qquad \begin{array}{c}\vdots\\ \varphi\end{array}}{\psi}\,{\to}E \tag{A.14}$$

we reduce to:

$$\begin{array}{c}\vdots\\ \varphi\\ \vdots\\ \psi.\end{array} \tag{A.15}$$

When none of these reductions can be applied anymore, we have a derivation where no main premise in the elimination rule is the conclusion of an introduction rule. It can, for instance, look like this:

$$\dfrac{\dfrac{\varphi \vee \varphi \qquad \dfrac{[\varphi] \qquad [\varphi]}{\varphi \wedge \varphi}\,{\wedge}I \qquad \dfrac{[\varphi] \qquad [\varphi]}{\varphi \wedge \varphi}\,{\wedge}I}{\varphi \wedge \varphi}\,{\vee}E}{\varphi}\,{\wedge}E \tag{A.16}$$

In this example, the conclusion of the or-elimination is still the main premise in the and-elimination. The derivation is thus still not normal according to our definition. The way we fix this does not diminish the derivation, but makes it bigger. We simply move the and-elimination up to the side premise in the or-elimination. This is called a *permutation*. In this way, we get the derivation:

$$\dfrac{\varphi \vee \varphi \qquad \dfrac{\dfrac{[\varphi] \qquad [\varphi]}{\varphi \wedge \varphi}\,{\wedge}I}{\varphi}\,{\wedge}E \qquad \dfrac{\dfrac{[\varphi] \qquad [\varphi]}{\varphi \wedge \varphi}\,{\wedge}I}{\varphi}\,{\wedge}E}{\varphi}\,{\vee}E \tag{A.17}$$

which in turn can be reduced to:

$$\dfrac{\varphi \vee \varphi \qquad [\varphi] \qquad [\varphi]}{\varphi}\,{\vee}E \tag{A.18}$$

More generally, when the conclusion in an $\vee E$ is the main premise in an elimination rule, we always move up the elimination rule (in two copies) to the side premise.

For our nullary disjunction $\bot$ we can do similar permutations. In $\bot E$, however, we have 0 side premises, so the permutation means that the elimination below is copied 0 times – that is, it disappears. For instance, we transform:

$$\dfrac{\dfrac{\begin{array}{c}\vdots\\ \bot\end{array}}{\varphi \wedge \psi}\,{\bot}E}{\varphi}\,{\wedge}E \tag{A.19}$$

into

$$\begin{array}{c} \vdots \\ \bot \\ \hline \varphi \end{array} \bot E \tag{A.20}$$

When $\bot E$ is followed by an elimination rule, we will always use such permutation.

We have now showed how every deviation from normality can be straightened. Since some of the transformations give smaller derivations, while some give larger ones, it is however not obvious that the process ends with a normal derivation. We shall not prove that *every* process ends in this way, but only that it is *possible* to get a normal derivation by applying the transformations above in a *certain ordering*. This is called *weak normalization*. A stronger result, *strong normalization* asserts that we would eventually get a normal derivation independently of the order in which we apply the transformations. This is more difficult to show, and it is not something we will need.

Since the difficulty in seeing immediately that the process ends is that the size of the derivations is not a measure which always decreases, we replace that by a better measure. We therefore need some notions.

▶ **A.21 Definition.** A main premise in an elimination rule is a *cut* if it is the conclusion in some rule which is not $\wedge E$ or $\rightarrow E$. A side premise in the rule $\vee E$ is a cut if the conclusion of the rule is a cut.

Our definition is somewhat simplified compared to the usual one. It works well in this context and it is easier to remember

A derivation is thus *normal* precisely when it does not contain any cuts. A derivation which is not normal is called *non-normal*.

▶ **A.22 Definition.** A *maximal cut* is a cut in the derivation such that no other cut in it contains more logical operations (we defined the number of operations in Exercise 6.3.3).

We shall now check that we can go through the normalization process so that it is guaranteed to finish and reach the promised result. It is sufficient to study RAA-free derivations. According to Glivenko's theorem, one can do without RAA except possibly in the last step, but in that the rest is an RAA-free derivation.

We need two measures for the proof: the number a of logical operations in maximal cuts and the number n of maximal cuts in the derivations. In fact, we will prove the theorem:

> For all natural numbers a, n it is true that if a derivation has n maximal cuts with a logical operations in each of them, the derivation can be normalized.

We will prove the theorem by a double induction on the natural numbers. In the proof, we therefore have access to the following two inductive hypothesis:

1. If a derivation has a maximal cut with less than a logical operations, it can be normalized.

2. If a derivation has less than n maximal cuts with a logical operations, it can be normalized.

Consider thus an RAA-free normal derivation with n maximal cuts and a logical operations in each. We will show that it can be normalized. We will do this by finding a suitable maximal cut in the derivation and remove it according to the transformations we have gone through. Afterwards, we will show that the resulting derivation can be normalized according to the inductive hypothesis. For this idea to succeed, we consider a maximal cut which does not have any other maximal cut underneath it in the derivation. We shall see that if such a cut is removed, the inductive hypothesis can be applied.

Such a cut cannot be a side premise in $\vee E$, because then the conclusion in the same rule is also a maximal cut, and we have just assumed that we are working with a maximal cut that does not have any other maximal cut underneath. The cut must, therefore, be a main premise in an elimination rule. That it is a cut means that it is the conclusion in an introduction rule, or in $\vee E$ or in $\bot E$. If it is a conclusion in an introduction rule, we can use the transformation rules (A.10)–(A.15). Since these eliminate a maximal cut, and no cut with the same or greater number of logical operations is created, the inductive hypothesis implies that the resulting derivation can be normalized (one may need to reduce the side derivations first to guarantee that the number of maximal cuts has not increased).

We are only left with the task of handling the cases in which the maximal cut is the conclusion in $\vee E$ or $\bot E$. The last case is as trivial as the one we just considered, so we are only left with considering $\vee E$. We assume, thus, that we have a derivation that looks as follows, where R denotes an elimination rule of which σ is the main premise and that has, possibly, side derivations:

$$\dfrac{\dfrac{\begin{matrix}\vdots\\ \varphi\vee\psi\end{matrix}\quad \begin{matrix}[\varphi]\\ \vdots\\ \sigma\end{matrix}\quad \begin{matrix}[\psi]\\ \vdots\\ \sigma\end{matrix}}{\sigma}\,{\scriptstyle\vee E}\qquad \vdots}{\tau}\,R \tag{A.23}$$

According to the inductive hypotheses, possible side derivations of R can be normalized, so we can assume they are normal and that the derivation contains, in all, n maximal cuts with a logical operations each. We transform according to the permutation rule and get the derivation:

$$\dfrac{\begin{matrix}\vdots\\ \varphi\vee\psi\end{matrix}\quad \dfrac{\begin{matrix}[\varphi]\\ \vdots\\ \sigma\end{matrix}\quad \vdots}{\tau}\,R\quad \dfrac{\begin{matrix}[\psi]\\ \vdots\\ \sigma\end{matrix}\quad \vdots}{\tau}\,R}{\tau}\,{\scriptstyle\vee E} \tag{A.24}$$

We must now show that this derivation can be normalized. If τ contains a smaller number of logical operations than σ then the new derivation has less than n cuts with a logical operations each. Every such cut must in fact come from a corresponding cut in the old derivation, and at least one has disappeared in the transformation. If, on the other hand, τ contains at least as many logical operations as σ, then τ could not have been a cut in the original derivation (since σ did not have a maximal cut underneath). Then, τ cannot be either a cut in the new derivation.

We have now gone through all possible ways a formula can be a cut and shown how these can be handled. Therefore we know that we can normalize every possible derivation. In addition, we have seen that in the part of the proof dealing with RAA-free derivations, we only used transformations which never added a rule that was not used previously. Hence, we know that when we normalize RAA-free derivations, we will never use any new rules. □

A.25 Example. Normalize

$$\dfrac{\dfrac{\varphi\vee\psi \quad \dfrac{[\varphi]\quad[\varphi]}{\varphi\wedge\varphi}\,{\scriptstyle\wedge I} \quad \dfrac{\dfrac{\neg\psi\quad[\psi]}{\bot}\,{\scriptstyle\to E}}{\varphi\wedge\varphi}\,{\scriptstyle\bot E}}{\varphi\wedge\varphi}\,{\scriptstyle\vee E}}{\varphi}\,{\scriptstyle\wedge E}$$

Solution. We use the transformation rules we have introduced and get:

$$\dfrac{\varphi \lor \psi \qquad [\varphi] \qquad \dfrac{\dfrac{\neg\psi \qquad [\psi]}{\bot}\,{\to}E}{\varphi}\,\bot E}{\varphi}\,\lor E$$

□

Reminder: by *normalize* we mean transform only by using the transformation rules we have introduced, in such a way that the end result is a normal derivation.

A.26 Exercise. Normalize

$$\dfrac{\dfrac{\dfrac{\dfrac{\bot}{\varphi \lor \bot}\,\lor I \qquad \dfrac{\neg\varphi \qquad [\varphi]}{\bot}\,{\to}E \qquad [\bot]}{\bot}\,\lor E}{\varphi \land \varphi}\,\bot E}{\varphi}\,\land E$$

We can start in three different ways. Try all three!

> *Hint.* Whatever we do, we end with a derivation with only one rule: $\bot E$.

A.27 Exercise. Normalize

$$\dfrac{\begin{array}{c}\vdots\\ \varphi \lor \neg\varphi\end{array} \qquad \dfrac{\varphi \to \psi \qquad [\varphi]}{\psi}\,{\to}E \qquad \dfrac{\neg\varphi \to \psi \qquad [\neg\varphi]}{\psi}\,{\to}E}{\psi}\,\lor E \tag{A.28}$$

where the dotted part is as in Example 5.4.3 (page 41).

> *Hint.* The result should be as in Exercise 7.3.25.

Solutions to the exercises

Solutions to old exams can be found at `http://www.math.su.se/`.

1.1.5 The axioms (comm), (ass), (id). Additionally, the right (abs), (distr) and left (inv).

1.1.6 $a \wedge a \overset{(\text{id})}{=} (a \vee 0) \wedge (a \vee 0) \overset{(\text{distr})}{=} a \vee (0 \wedge 0) \overset{(\text{abs})}{=} a \vee 0 \overset{(\text{id})}{=} a$

1.1.7 $\neg\neg a \overset{(\text{id})}{=} \neg\neg a \vee 0 \overset{(\text{inv})}{=} \neg\neg a \vee (a \wedge \neg a) \overset{(\text{distr})}{=} (\neg\neg a \vee a) \wedge (\neg\neg a \vee \neg a) \overset{(\text{comm})}{=} (\neg\neg a \vee a) \wedge (\neg a \vee \neg\neg a) \overset{(\text{inv})}{=} (\neg\neg a \vee a) \wedge 1 \overset{(\text{id})}{=} \neg\neg a \vee a$

In the same way, one later shows that $a = a \vee \neg\neg a$: $a \overset{(\text{id})}{=} a \vee 0 \overset{(\text{inv})}{=} a \vee (\neg a \wedge \neg\neg a) \overset{(\text{distr})}{=} (a \vee \neg a) \wedge (a \vee \neg\neg a) \overset{(\text{inv})}{=} 1 \wedge (a \vee \neg\neg a) \overset{(\text{comm})}{=} (a \vee \neg\neg a) \wedge 1 \overset{(\text{id})}{=} a \vee \neg\neg a$

1.2.5 Because of (id) and (abs), it is clear how the table entries under $\wedge$ and $\vee$ should be. Because of this and (inv), the table entries under $\neg$ is determined as well.

1.3.2 $\neg 1 \overset{(\text{id})}{=} \neg 1 \wedge 1 \overset{(\text{comm})}{=} 1 \wedge \neg 1 \overset{(\text{inv})}{=} 0$

1.3.4 $a \wedge (a \vee b) \overset{(\text{id})}{=} (a \vee 0) \wedge (a \vee b) \overset{(\text{distr})}{=} a \vee (0 \wedge b) \overset{(\text{comm})}{=} a \vee (b \wedge 0) \overset{(\text{abs})}{=} a \vee 0 \overset{(\text{id})}{=} a$

1.3.6 $b \overset{(\text{id})}{=} b \wedge 1 \overset{(1.3.1)}{=} b \wedge \neg 0 = b \wedge \neg(a \vee b) \overset{(\text{dM})}{=} b \wedge (\neg a \wedge \neg b) \overset{(\text{comm})}{=} b \wedge (\neg b \wedge \neg a) \overset{(\text{ass})}{=} (b \wedge \neg b) \wedge \neg a \overset{(\text{inv})}{=} 0 \wedge \neg a \overset{(\text{comm})}{=} \neg a \wedge 0 \overset{(\text{abs})}{=} 0$

1.3.8 Assume that $a \wedge b = 0$. In the Boolean algebra with two elements we have either $b = 0$ or $b = 1$. If $b = 0$ then we are done. Otherwise we have $b = 1$, and then $a = a \wedge 1 = a \wedge b = 0$.

In algebras with more than two elements we cannot argue in that way. We have seen in Example 1.2.6 that we can have $s \wedge t = 0$ without having $s = 0$ nor $t = 0$.

1.3.10 In a Boolean algebra. $0 \leq 1$ means that $0 \wedge 1 = 0$, according to Definition 1.3.9. This is true because of axiom (id).

1.3.11 By $a \wedge b \leq b$ it is meant $(a \wedge b) \wedge b = a \wedge b$, according to Definition 1.3.9. This is shown using (ass) and (idemp).

1.3.12 Reflexivity: $a \leq a$ means that $a \wedge a = a$, which is an axiom (idemp).

Transitivity: suppose $a \leq b$ and $b \leq c$; that is, $a \wedge b = a$ and $b \wedge c = b$. Show that $a \leq c$; that is, $a \wedge c = a$. $a \wedge c = (a \wedge b) \wedge c \overset{(\text{ass})}{=} a \wedge (b \wedge c) = a \wedge b = a$.

Antisymmetry: suppose $a \wedge b = a$ and $b \wedge a = b$. Then $a = a \wedge b \overset{(\text{comm})}{=} b \wedge a = b$.

1.3.13 $a \leq (a \vee b)$ means that $a \wedge (a \vee b) = a$, which is an absorption rule.

$b \leq (a \vee b)$ means that $b \wedge (a \vee b) = b$. To prove this, we first apply (comm), and then an absorption rule: $b \wedge (a \vee b) \overset{(\text{comm})}{=} b \wedge (b \vee a) \overset{(\text{abs})}{=} b$.

Suppose now that $a \leq c$ and $b \leq c$; that is, $a \wedge c = a$ and $b \wedge c = b$. We then have $(a \vee b) \wedge c = (a \wedge c) \vee (b \wedge c) = a \vee b$.

1.3.14 This means that $(a \wedge b) \leq a$, $(a \wedge b) \leq b$, and: if $c \leq a$ and $c \leq b$, then $c \leq (a \wedge b)$. This is proven in a completely analogous way to the previous exercise.

1.3.15 In the Boolean algebra with two elements, 1 is an atom, since $1 \neq 0$ and if $c \leq 1$ for some $c \neq 0$, then $c = 1$, since it is the only element which is not 0.

In the algebra of subsets of $\{1, 2, 3\}$ we have the atoms $\{1\}$, $\{2\}$, $\{3\}$. To show that $\{1\}$ is an atom, we assert that $\{1\} \neq \emptyset$ and that if $M \subseteq \{1\}$ and $M \neq \emptyset$ then M has to contain at least one element; hence $1 \in M$, and thus $M = \{1\}$. Similarly we show that $\{2\}$ and $\{3\}$ are atoms.

1.3.16 Assume that $a \leq b$; that is, $a \wedge b = a$. We will show that $(a \vee c) \leq (b \vee c)$; that is, $(a \vee c) \wedge (b \vee c) = (a \vee c)$. The left term can be rewritten using (distr) as $(a \wedge b) \vee c$, which is equal to $a \vee c$.

1.3.17 Assume that $a \leq b$; that is, $a \wedge b = a$. We will show that $\neg b \leq \neg a$; that is, $\neg b \wedge \neg a = \neg b$:

$$\neg b \wedge \neg a = \neg b \wedge \neg(a \wedge b) \overset{\text{(dM)}}{=} \neg b \wedge (\neg a \vee \neg b) \overset{\text{(abs)}}{=} \neg b.$$

1.4.4 $x \vee y \wedge y \vee \neg x = x \vee (y \wedge y) \vee \neg x = 1$.

$x \wedge y \vee y \wedge \neg x = (x \wedge y) \vee (y \wedge \neg x) = y \wedge (x \vee \neg x) = y \wedge 1 = y$.

$\neg(\neg(x \wedge y) \vee x) \vee y = (x \wedge y \wedge \neg x) \vee y = 0 \vee y = y$.

1.5.8 Expressions 0, $x \wedge y \wedge z$, x, $x \vee \neg x$ are in disjunctive normal form.

Expressions 0, $(x \vee y) \wedge z$, $x \wedge y \wedge z$, x are in conjunctive normal form.

It is not possible to decide in which form $a \vee b$ is. If a and b denote Boolean expressions, it depends on these expresions whether $a \vee b$ is in disjunctive or conjunctive normal form.

1.5.9 $x \wedge y \vee x \wedge z$, $x \wedge \neg y \wedge z$, $\neg y \wedge \neg z$.

1.6.6 Write the left hand side in disjunctive normal form. One gets then the equation $x \wedge \neg y \wedge z = 0$. We do not get any further than this.

1.6.7 Write the left hand side in disjunctive normal form. One gets then the equation $\neg y \wedge \neg z = 0$. We do not get any further than this.

2.1.35 Replace by the equivalent equation $x \wedge y \wedge \neg z = 0$. Better than that one cannot answer, in general. In the two elements algebra one can pick out the solution explicitly: all the values of x, y, z except for $(x, y, z) = (1, 1, 0)$ are solutions.

2.1.36 We first simplify the left hand side to $x \wedge \neg y \wedge z = \neg y \wedge \neg z$. It splits into the equalities $x \wedge \neg y \wedge z \leq \neg y \wedge \neg z$ respectively $\neg y \wedge \neg z \leq x \wedge \neg y \wedge z$. Consider the second one. It can be split into three different equalities (with the right hand sides x, $\neg y$, z). The inequality which has z on the right hand side ($\neg y \wedge \neg z \leq z$) can be rewritten as the equation $\neg y \wedge \neg z \wedge \neg z = 0$, that is $\neg y \wedge \neg z = 0$. This tells us nothing about the right hand side in the original equation! The original equation is thus equivalent to $x \wedge \neg y \wedge z = 0$, $\neg y \wedge \neg z = 0$.

2.1.37 Look first at the solution of the previous exercise to see how to change the first equation into the equivalent equation system: $x \wedge \neg y \wedge z = 0$, $\neg y \wedge \neg z = 0$. The original system can thus (as the inequality can be rewritten as an equation) be written as $x\wedge\neg y\wedge z = 0$, $\neg y\wedge\neg z = 0$, $x\wedge y\wedge\neg z = 0$, $y\wedge z = 0$. If one colours the corresponding areas in a Venn diagram, one sees that the whole area corresponding to x is coloured, so one can wonder whether $x = 0$ is possible to be derived from the first four equations. This is in fact possible, but you can be satisfied if you have come this far. If you want to derive $x = 0$ you can observe that $x = (x\wedge y\wedge z)\vee(x\wedge y\wedge\neg z)\vee(x\wedge\neg y\wedge z)\vee(x\wedge\neg y\wedge\neg z) = 0\vee 0\vee 0\vee 0 = 0$. Hence, the given system is equivalent to $x = 0$, $\neg y \wedge \neg z = 0$, $y \wedge z = 0$. The last two equations can be written as $\neg z \leq y$ respectively $y \leq \neg z$, so together they give $y = \neg z$. The given equation system is thus equivalent to the system $x = 0,\ y = \neg z$.

If while doing the previous work one notices that one has both $\neg y \wedge \neg z = 0$ and $y \wedge z = 0$, and that together this gives $y = \neg z$, then one can replace, in the original system, all occurrences of y with $\neg z$, and simplify. Then we get the system $x \wedge z = 0$, $x \wedge \neg z = 0$, $y = \neg z$. From this one gets $x = x \wedge (z \vee \neg z) = x\wedge z\vee x\wedge\neg z = 0\vee 0 = 0$. Afterwards, we arrive at the solution $x = 0,\ y = \neg z$. But, as we said, this requires to pay close attention. The first solution was more routine.

2.2.4 We will have $(1 \wedge b) \leq c \iff 1 \leq (b \to c)$; that is, $b \to c$ should be 1 precisely when $b \leq c$, i.e., in every case except when $b = 1$ and $c = 0$.

2.2.12 $a \to 0 = \neg a \vee 0 = \neg a$.

2.2.15 In the rows where $a = b$ we get 1, while in the rows where $a \neq b$ we get 0.

2.2.17 $(\neg x \wedge y \wedge \neg z) \vee (\neg x \wedge y \wedge z) \vee (x \wedge \neg y \wedge \neg z) \vee (x \wedge y \wedge z)$

2.2.18 $x \wedge (x \to y) = x \wedge (\neg x \vee y) = x \wedge \neg x \vee x \wedge y = x \wedge y$.

$\neg x \to x = \neg\neg x \vee x = x$.

$(x \wedge \neg x) \to y = 0 \to y = 1$.

$x\vee y\to\neg x\wedge y = \neg(x\vee y)\vee(\neg x\wedge y) = (\neg x\wedge\neg y)\vee(\neg x\wedge y) = \neg x\wedge(\neg y\vee y) = \neg x$.

2.2.19 The equation $y = \neg x$ is equivalent to the two inequalities $y \leq \neg x$ and $\neg x \leq y$, which are in turn equivalent to the two equations $y \wedge x = 0$ and $\neg x \wedge \neg y = 0$. If the second of these equations is negated, we get $x \vee y = 1$.

2.2.22 If the right hand side is written in disjunctive normal form, it becomes $\neg x\vee\neg y\vee(x\wedge z)$. Since $\neg x \leq HL = VL \leq x$ we have $\neg x \leq x$, thus $\neg x\wedge\neg x = 0$, so that $x = 1$. When we put this into the original equation, it simplifies to $y\to z = y\to z$, which is true for all replacements. The solutions of the equation are then $x = 1$, while y and z are arbitrary. We can come to the same conclusion using the standard methods, though it can take a little bit longer.

3.2.28 $1 + 2 = 1 + s(1) = s(1 + 1) = s(1 + s(0)) + s(s(1 + 0)) = s(s(1))$.

$1\cdot 2 = s(0)\cdot 2 = 0\cdot 2+2 = 0+s(1) = s(0+1) = s(0+s(0)) = s(s(0+0)) = s(s(0))$

$1-2 = 1-s(1) = p(1-1) = p(1-s(0)) = p(p(1-0)) = p(p(1)) = p(p(s(0))) = p(0) = 0$

3.2.29 $f(a, b) = \max(a, b)$.

3.3.3 $a \to 0 = \neg a \vee 0 = \neg a$.

$a \to 1 = \neg a \vee 1 = 1$.

4.2.8 $[\![\neg\varphi]\!] = [\![\varphi \to \bot]\!] = [\![\varphi]\!] \to [\![\bot]\!] = [\![\varphi]\!] \to 0 = \neg[\![\varphi]\!]$

$[\![\varphi\leftrightarrow\psi]\!] = [\![(\varphi\to\psi)\wedge(\psi\to\varphi)]\!] = [\![\varphi\to\psi]\!]\wedge[\![\psi\to\varphi]\!] = ([\![\varphi]\!]\to[\![\psi]\!])\wedge([\![\psi]\!]\to[\![\varphi]\!]) =$

$[\![\varphi]\!] \leftrightarrow [\![\psi]\!]$

4.2.17 $[\![\neg(P_2 \to \neg P_3) \wedge (P_1 \to P_5)]\!] = [\![\neg(P_2 \to \neg P_3)]\!] \wedge [\![(P_1 \to P_5)]\!] = \neg[\![P_2 \to \neg P_3]\!] \wedge ([\![P_1]\!] \to [\![P_5]\!]) = \neg([\![P_2]\!] \to [\![\neg P_3]\!]) \wedge (0 \to 1) = \neg(0 \to \neg[\![P_3]\!]) \wedge (0 \to 1) = \neg 1 \wedge 1 = 0.$

4.2.31 Assume that $\varphi \leftrightarrow \psi$ is a tautology; that is, its truth value is 1 in every interpretation. If one considers the truth table for $\leftrightarrow$ one sees that φ and ψ have the same truth values in all interpretations. The converse is also evident from the truth table.

4.2.32 $\neg(P_1 \wedge P_2) \leftrightarrow (P_1 \to \neg P_2)$ is a tautology if and only if $\neg(P_1 \wedge P_2) \approx (P_1 \to \neg P_2)$. Since $P_1 \to \neg P_2 \approx \neg P_1 \vee \neg P_2$ the answer is yes according to de Morgan's laws.

$(P_1 \to P_2) \vee (P_2 \to P_3) \approx \neg P_1 \vee P_2 \vee \neg P_2 \vee P_3 \approx \top$. This is a tautology.

$(P_1 \to (P_2 \to P_3)) \leftrightarrow ((P_1 \wedge P_2) \to P_3)$ is a tautology if and only if $(P_1 \to (P_2 \to P_3)) \approx ((P_1 \wedge P_2) \to P_3)$. The left hand side simplifies to $(P_1 \to (P_2 \to P_3)) \approx \neg P_1 \vee \neg P_2 \vee P_3$ and the right hand side simplifies to $(P_1 \wedge P_2) \to P_3 \approx \neg(P_1 \wedge P_2) \vee P_3 \approx \neg P_1 \vee \neg P_2 \vee P_3$. Since these terms simplify to the same expression, this is a tautology.

$((P_1 \wedge P_4) \to (P_2 \vee P_3)) \leftrightarrow (\neg P_1 \vee P_2 \vee P_3 \vee P_4)$ is a tautology if and only if $((P_1 \wedge P_4) \to (P_2 \vee P_3)) \approx (\neg P_1 \vee P_2 \vee P_3 \vee P_4)$. We simplify the left hand side as: $(P_1 \wedge P_4) \to (P_2 \vee P_3) \approx \neg(P_1 \wedge P_4) \vee P_2 \vee P_3 \approx \neg P_1 \vee \neg P_4 \vee P_2 \vee P_3$. This is different from the right hand side: if P_1 is interpreted as a true proposition while the rest of the variables are interpreted as false propositions, then the left hand side is true while the right hand side is false. Answer: this formula is not a tautology.

4.2.41 In the case $n = 1$ one gets that $\varphi_1 \vDash \varphi$ means that φ is true in every interpretation where φ_1 is true.

In the case $n = 0$ one gets that $\vDash \varphi$ means that φ is true in every interpretation; that is, φ is tautology. Indeed, one should check that φ is true in every interpretation in which all the formulas on the left hand side of $\vDash$ are true, but since there are no formulas there, every interpretation satisfies this criteria. In conclusion: $\vDash \varphi$ is a way of writing that φ is a tautology.

5.3.6 It is enough to apply one rule: $\vee I$.

5.3.7 End the derivation by using $\vee E$. Above this line put $\varphi \vee \psi$ and derivations from φ respectively ψ to $\psi \vee \varphi$ (use $\vee I$).

5.3.8 End by using $\vee E$. Above the line put $\varphi \vee \bot$ and two derivations: from φ, respectively $\bot$, to φ. The first consists of the formula φ itself. The other consist of a single application of $\bot E$.

5.3.9 End with $\to I$. Above it, use $\vee E$.

5.3.10 End with $\to I$. Above it, use $\vee E$.

5.4.4 A formula of the form $\neg\psi$. If every assumption is discharged, the last rule has to discharge ψ. Thus, one derives $\neg\neg(\varphi \vee \neg\varphi)$.

6.1.18 As (6.1.1), respectively (6.1.2).

6.1.22 The formula can be derived using $\to I$ twice. Thus, it is a tautology according to the soundness theorem.

6.1.25 If one could derive it, according to the soundness theorem it would be a tautology. But it is false in the interpretation where P_1 is false and P_2 is true.

6.1.28 The first subset is consistent since there is a model of it. If one could derive $\bot$ from it, then, according to the soundness theorem, $\bot$ should be true in all interpretations for which the formulas in the set are true, but in the model of the subset, $\bot$ is still false.

The other is inconsistent. This is more easily seen by constructing a derivation of $\bot$ from it. Start by showing that $P_4 \to P_2, P_2 \to P_3, P_3 \to \neg P_4 \vdash \neg P_4$ and that $P_1 \to P_2, P_2 \to P_3, P_3 \to \neg P_1, \neg P_4 \to P_1 \vdash P_4$. Then join together the two derivations into one derivation of $\bot$ and check that all undischarged assumptions are in the set $\{P_1 \to P_2, P_2 \to P_3, P_3 \to \neg P_1, P_4 \to P_2, P_3 \to \neg P_4, \neg P_4 \to P_1\}$.

6.1.29 Let, for instance, $\varphi = \top$ and $\psi = \bot$. Then the formula is false, and cannot, according to the soundness theorem, be derived.

6.1.31 In that case the last but one formula in the derivation would either be P_1 or $P_2 \vee P_3$, but none of these can be derived from $(P_1 \vee P_2) \vee P_3$. Indeed, if one could derive P_1 from $(P_1 \vee P_2) \vee P_3$ then, according to the soundness theorem, P_1 would be true in all interpretations in which $(P_1 \vee P_2) \vee P_3$ is true. But if P_2 is true, then $(P_1 \vee P_2) \vee P_3$ is true even when P_1 is false. If, instead, $P_2 \vee P_3$ could be derived from $(P_1 \vee P_2) \vee P_3$ then, according to the soundness theorem, $P_2 \vee P_3$ should be true in all interpretations in which $(P_1 \vee P_2) \vee P_3$ is tue. But if P_1 is true, then $(P_1 \vee P_2) \vee P_3$ is true even if $P_2 \vee P_3$ is not.

6.1.32 It is not possible to have this for every choice, as we have seen in Exercise 6.1.31. But it is possible if $\varphi = \top$, in which case we could end with $\vee I$ from $\top$.

6.1.33 In this case we should have a derivation of $\bot$ by removing the last step of the given derivation. But then, according to the soundness theorem, $\bot$ should be a tautology, which is not the case.

6.1.34 Otherwise, if one removes the last step of such a derivation, one should have a derivation from $\varphi \vee \psi$ to either φ or ψ. But since φ and ψ stand for arbitrary formulas, we could, for instance, insert P_1 forφ and P_2 for ψ and get thus a derivation of either P_1 or P_2 from $P_1 \vee P_2$. None of these is possible, according to the soundness theorem, as neither P_1 nor P_2 should necessarily be true just because $P_1 \vee P_2$ is (though at least one of them should be true).

6.1.35 Assume that we had a derivation of the formula ending in two introduction rules. Then the last one should be a $\to I$ and the last but one a $\vee I$. But then we should have the task of deriving either P_1 or P_2 (depending on which of the $\vee I$-rules we chose) from $P_1 \vee P_2$. We have seen in the solutions of previous exercises that this is impossible. But one can derive $(P_1 \vee P_2) \to (P_1 \vee P_2)$ using only one $\to I$-rule.

6.1.36 Assume that one can end with an introduction rule. Then we would have, above that point, the task of deriving P_1 from P_1. It is impossible, according to the soundness theorem, to derive P_1 without any discharged assumption, so P_1 must indeed be used as an assumption. But if we do not discharge any assumptions in the whole derivation, P_1 would be left as an undischarged assumption. This contradicts the fact that we have "derived $P_1 \to P_1$", since by that we mean that we have created a derivation without any undischarged assumptions.

6.1.37 That $\varphi_1, \ldots, \varphi_n \vDash \varphi$ means that in every interpretation where $\varphi_1, \ldots, \varphi_n$ are true, φ is also true. But $\varphi_1, \ldots, \varphi_n$ are true if and only if $[\![\varphi_1]\!] \wedge \cdots \wedge [\![\varphi_n]\!] = 1$. Thus it follows from $\varphi_1, \ldots, \varphi_n \vDash \varphi$ that if $[\![\varphi_1]\!] \wedge \cdots \wedge [\![\varphi_n]\!] = 1$ then $[\![\varphi]\!] = 1$; that is, $[\![\varphi_1]\!] \wedge \cdots \wedge [\![\varphi_n]\!] \leq [\![\varphi]\!]$. The converse is shown by following this argument backwards.

Then we can show that $\varphi_1, \ldots, \varphi_n, \varphi \vDash \psi \iff \varphi_1, \ldots, \varphi_n \vDash \varphi \to \psi$ is clearly equivalent to having the following condition: $[\![\varphi_1]\!] \wedge \cdots \wedge [\![\varphi_n]\!] \wedge [\![\varphi]\!] \leq [\![\psi]\!]$ is true in every interpretation if and only if $[\![\varphi_1]\!] \wedge \cdots \wedge [\![\varphi_n]\!] \leq [\![\varphi]\!] \to [\![\psi]\!]$ is true in every interpretation. But this condition follows directly from the Galois connection. (This shows that $\to$ really is what in the "language" corresponds to $\vDash$, which is a relation we do not have in our set of formulas.)

6.3.1 We have $1 \wedge a \leq b$ if and only if $1 \leq a \to b$, according to the Galois connection.

6.3.2 $f(a, b) = a + b$.

6.3.3 $a(\top \leftrightarrow \neg P_1) = a((\top \to (P_1 \to \bot)) \wedge ((P_1 \to \bot) \to \top)) = 9$.

6.3.4 $(\bot \wedge \bot) \neq \bot$, since they are different formulas. However they have the same truth values, so $(\bot \wedge \bot) \approx \bot$. $(P_1 \leftrightarrow \bot) = (\neg P_1 \wedge (\bot \to P_1))$ is true, because of how $\neg$ and $\leftrightarrow$ are defined. Therefore, we also have $(P_1 \leftrightarrow \bot) \approx (\neg P_1 \wedge (\bot \to P_1))$.

6.3.5 $\varphi = P_1 \wedge (P_2 \vee P_3) \wedge \neg(P_2 \wedge P_3)$.

7.1.4 All formulas which are on a line are premises. In $\vee E$ the two premises to the right are side premises. In $\to E$ the premise on the right is a side premise.

7.3.2 The main premise in the last rule is $\bot$. According to the theorem above, this is a subformula of some undischarged assumption.

7.3.3 Normalize a given derivation of φ. According to the previous theorem, this cannot end in an elimination rule. If it did, its main premise would be a subformula of some undischarged assumption, but no such thing exists.

7.3.4 No. If one had such a derivation, one could normalize it and get a normal derivation without undischarged assumptions in which only the rules $\vee I$, $\to I$, $\to E$ and $\wedge E$ are used. Such a thing cannot end in an elimination rule, and because of the form of the formula (it is an implication formula, i.e., the outmost connective is an implication) the last rule has to be $\to I$. The question is now whether one can derive $\neg P_1 \vee \neg P_2$ from $\neg(P_1 \wedge P_2)$. The last step in this derivation cannot be $\vee I$ (according to the soundness theorem), nor $\to I$ (since $\neg P_1 \vee \neg P_2$ is not an implication formula), nor $\to E$ (since the main premise must contain $\neg P_1 \vee \neg P_2$ as subformula, and in turn, according to the previous theorem, has to be also a subformula in $\neg(P_1 \wedge P_2)$) and nor $\wedge E$ (for the same reason). Hence we cannot continue using only these four rules.

7.3.7 Assume $\vdash P_1$. In this case there is a normal derivation of P_1 without any undischarged assumptions. Assume first that it does not contain RAA. Then it can only contain subformulas of P_1, which is impossible, as no rules can be applied to these. That is, a possible derivation has to use RAA. But then there is, according to Glivenko's theorem, a derivation which uses RAA in the last step, and above this last step there is a normal derivation from $\neg P_1$ to $\bot$. The only subformula of these, other than the formulas themselves, is P_1, but then the only rules we can use are $\to I$, $\to E$ and $\bot E$. One is then forced to go around in circles when seeking a normal derivation with only these rules.

7.3.8 Assume that $\vdash \neg P_1$. In this case there exists a normal derivation of $\neg P_1$ without any undischarged assumptions. Assume first that it does not contain RAA. Then it can only contain subformulas of $\neg P_1$. It cannot end with an elimination rule, so it has to end with $\to I$. Above it there should be a normal derivation from P_1 to $\bot$, but the only subformulas of these are the formulas themselves, so no rules except $\bot E$ are possible. Not even this one can be used, acording to Exercise 7.3.2. Therefore, every derivation of $\neg P_1$ must contain RAA, but according to Glivenko's theorem, there would exist, in this case, a derivation which ends in RAA and does not have RAA anywhere else. Above RAA we must have a normal derivation from $\neg\neg P_1$ to $\bot$. For the same reasons as in the previous exercise, one can see that this is not possible.

7.3.9 Assume that you have a derivation without RAA. By normalizing it, one could get a normal derivation without RAA. According to Theorem 7.3.1 it cannot end with an elimination rule, so the last rule must be $\vee I$. Above it, we have a normal derivation of P_1 or of $\neg P_1$, but you have shown in the last couple of exercises that this is impossible.

7.3.10 No. If one had such a derivation, one should be able to normalize it to get a RAA-free normal derivation without undischarged assumptions. Let us consider such a derivation. It cannot end in an elimination rule, so it has to end with $\to I$. The step above it cannot be an introduction rule, since $\vee I$ is the only candidate, but its premises cannot be derived from $\neg(P_1 \wedge P_2)$ (which can be shown using both the soundness theorem and the subformula property). Thus, it has to be an elimination rule, which is our next step. Then the main premise has to be a subformula of $\neg(P_1 \wedge P_2)$, which excludes the possibility of having $\vee E$. For the same reason $\wedge E$ and $\to E$ are excluded, since the main premise in such cases should also contain $\neg P_1 \vee \neg P_2$ as a subformula, which is impossible. We are only left with $\bot E$. But it is impossible to derive $\bot$ from $\neg(P_1 \wedge P_2)$, which can be shown by using the soundness theorem or the subformula property.

7.3.11 Because of the subformula property, one can only use rules which contain the operation $\vee$.

7.3.25 We first try to find a normal derivation without RAA. Since every formula in the derivation must be, in this case, a subformula of some undischarged assumption or of the conclusion P_2, it is only the formulas $\neg P_1$ and $\bot$ the ones we have to work with, in addition to the ones occurring in the exercise. In particular, it follows that we just need to investigate the rules $\to I$, $\to E$ and $\bot E$. We cannot end with an introduction rule, so we must end with an elimination rule, whose main premise is a subformula of some undischarged assumption; that is, of $P_1 \to P_2$ or $\neg P_1 \to P_2$ or $\neg P_1$ or $\bot$. If $\to E$ is the rule, the main premise has to be $P_1 \to P_2$ or $\neg P_1 \to P_2$, since it should contain the conclusion P_2. Then we must have P_1 respectively $\neg P_1$ as side premises, but these are not derivable. Therefore, this is not a viable path to take. In the same way, it is not viable to end with $\bot E$. We have therefore excluded the possibility of doing this derivation without RAA. According to Glivenko's theorem, we know that there is a derivation which concludes with RAA. We therefore go on with the problem of deriving $\bot$ from $P_1 \to P_2,\ \neg P_1 \to P_2,\ \neg P_2$, without using RAA. We look for a normal derivation. The last rule must therefore be $\to E$ with $\neg P_2$ as main premise. It remains to derive the side premise P_2. With the same reasoning as above, we conclude that the last rule should be $\to E$ with $\neg P_1 \to P_2$ as the main premise. The side premise $\neg P_1$ is derived from the assumption P_1 together with $P_1 \to P_2$ and $\neg P_2$.

8.1.3 If $\Gamma = \{\gamma_1, \ldots, \gamma_n\}$ was closed under derivations, $\gamma_1 \wedge \cdots \wedge \gamma_n \in \Gamma$, which is impossible, since it contains more logical operations than $\gamma_1, \ldots, \gamma_n$. (Note that in the case $\Gamma = \emptyset$ then the conjunction in this argument is $\top$.)

We can also show the result by noting that all of the following formulas are derivable, and hence necessarily included in any set closed under derivations: $\top, \top \wedge \top, (\top \wedge \top) \wedge \top, \ldots$.

8.1.13 It is sufficient to show that $\Gamma^* \cup \{\varphi\}$ is consistent. Assume therefore that it was inconsistent. Consider a derivation of $\bot$, without any undischarged assumptions, except, possibly, formulas in $\Gamma^* \cup \{\varphi\}$. Continue the derivation downwards with a $\to I$ and discharge all assumptions of φ. One then gets a derivation showing that $\Gamma^* \vdash \neg\varphi$. But this contradicts $\neg\varphi \notin \Gamma^*$ since Γ^* is closed under derivations.

8.1.14 Construct a derivation of φ from $\neg\psi$ and $\varphi \vee \psi$ by ending with $\vee E$ where $\varphi \vee \psi$ is the main premise. In the derivation on the right, one can apply that $\neg\psi$ together with ψ gives $\bot$ and concludes that φ with the help of $\bot E$.

8.1.15 Show $\psi \vdash \varphi \to \psi$ using $\to I$ and the fact that Γ^* is closed under derivations.

8.1.16 Use the fact that if $\varphi \notin \Gamma^*$, then $\neg\varphi \in \Gamma^*$ and that $\neg\varphi \vdash \varphi \to \psi$, along with the fact that Γ^* is closed under derivations.

8.1.17 That $\{P_1, P_2, P_3, \neg P_1 \vee \neg P_2\}$ is inconsistent is most easily shown by constructing a derivation. The set of all propositional variables is not maximally consistent, since it is not closed under derivations.

8.2.1 Assume that $\Gamma \vdash \bot$. It follows from the soundness theorem that $\Gamma \models \bot$. But then every model of Γ would be a model of $\bot$, which is impossible if there is a model of Γ, since nothing is a model of $\bot$.

8.2.4 Consider a formula which is true in all interpretation in which a certain set of formulas is true. Then there is a derivation of the formula in natural deduction, without any undischarged assumptions, except, possibly, those in the given set.

8.2.5 The first part of the exercise is to prove that $\vdash \varphi \iff \models \varphi$. This is just a special case of the soundness theorem and the completeness theorem. The second part is to prove that $\Gamma \nvdash \bot \iff \Gamma \not\models \bot$. It also follows immediately from the soundness theorem and the completeness theorem.

8.2.6 According to the previous exercise, we have $\vdash \varphi \leftrightarrow \psi$ if and only if $\varphi \leftrightarrow \psi$ is a tautology. The rest follows as in the previous exercise.

8.2.7 According to the previous exercise, it is enough to show that $\vdash (\varphi \vee \psi) \to (\varphi \wedge \psi)$ if and only if $\vdash \varphi \leftrightarrow \psi$. This is most easily shown by explaining how a derivation of one could be used to construct a derivation of the other.

8.2.8 The resulting formulas are of the form $\varphi \leftrightarrow \psi$. These are, according to the previous exercise, derivable if and only if $\varphi \approx \psi$. We need therefore to prove the latter. But that φ and ψ have the same value in every interpretation is guaranteed by the fact that they are in the left, respectively right hand side of the Boolean axioms, since the truth values are computed in Boolean algebras.

9.1.15 "x_2 occurs in x_{23}" means that $2 = 23$, which is not true; thus, the answer is no. On the other hand, the answer to the other questions are yes, with the only exception that x_2 does not occur in $f_3(x_0, f_1)$.

9.1.17 $f_3(x_0, f_1)[x_1/x_0] = f_3(x_1, f_1)$

$f_3(x_0, x_1)[x_1/x_0][x_0/x_1] = f_3(x_0, x_0)$

$f_3(x_0, f_1)[f_4(f_3(x_0, x_1), f_3(x_2, x_3))/x_2] = f_3(x_0, f_1)$

9.1.18 Give a proof by induction. Split t into the two possible cases: t can be a variable or a function symbol with arguments.

9.1.19 If $t = x_i \neq x_j$, then $t[x_i/x_j][x_j/x_i] = x_j \neq t$. To show that if x_i does not occur in t then $t[x_i/x_j][x_j/x_i] = t$, we use a proof by induction. If $t = x_k$, we consider two cases. If $k = j$ then we have $t[x_i/x_j][x_j/x_i] = x_j = t$. If $k \neq j$ then we have $i \neq k$, and thus $t[x_i/x_j][x_j/x_i] = x_k[x_i/x_j][x_j/x_i] = x_k[x_j/x_i]$. If x_i does not occur in t then we have $i \neq k$, so that $x_k[x_j/x_i] = x_k = t$.

If $t = f(\ldots)$ then the results follows immediately from the inductive hypothesis, since the substitution is done by substituting in every argument.

9.2.4 a) The tree looks precisely as in predicate logic.

b) Every rule from the definition of Form in propositional logic is also a rule in the new definition of Form.

c) x_0 and x_1, for instance.

d) $x_0 \doteq x_1$ is an example; another one is $\forall x_0 \top$.

9.2.6 a) According to the definition, it is propositional if P_1 and P_2 are propositional, which is the case by definition.

b) (With my examples:) "$x_0 \doteq x_1$ propositional" is false according to the first row of the definition. "$\forall x_0 \top$ propositional" is false according to the last row of the definition.

9.2.7 "x_i occurs in $t_1 \doteq t_2$" is defined as x_i occurs in t_1 or in t_2.

"x_i occurs in $\top$" is defined as false, and in the same way one deals with $\bot$.

"x_i occurs in $\varphi \wedge \psi$ is defined as x_i occurs in φ or in ψ – and in the same way it is defined for $\vee$ and $\to$.

"x_i occurs in $\forall x_j \varphi$" is defined as $i = j$ or x_i occurs in φ – and in the same way for $\exists$.

9.2.8 Induction over Form. For many sorts of formulas it is vacuously true that "if φ is a propositional formula, it is false that x_i occurs in φ". For the other cases it follows immediately from the induction step.

9.2.10 $(x_1 \doteq x_2 \wedge P_1(f_1(x_1, x_2)))[f_2/x_1] = (f_2 \doteq x_2 \wedge P_1(f_1(f_2, x_2)))$

$(x_1 \doteq x_2 \wedge \forall x_1(x_1 \doteq x_2))[f_2/x_1] = (f_2 \doteq x_2 \wedge \forall x_1(x_1 \doteq x_2))$

$\forall x_1 \forall x_2(x_1 \doteq x_2 \wedge x_2 \doteq x_3)[x_3/x_2] = \forall x_1 \forall x_2(x_1 \doteq x_3 \wedge x_3 \doteq x_3)$ (precedence rule: substitution binds stronger than quantifiers).

9.2.15 In atomic formulas, it is false that x_i occurs bound, and the same for $\top$ and $\bot$. In $\varphi \wedge \psi$, $\varphi \vee \psi$ and $\varphi \to \psi$ bound means that the variable occurs bound in φ or in ψ. In $\forall x_j \varphi$, x_i occurs bound if $i = j$ or x_i occurs bound in φ. Similarly for the case of $\exists$.

9.2.16 a) Yes. b) Yes. c) Yes. d) No. e) Yes. f) Yes. g) No. h) Yes ($\neg$ is defined as $\to\bot$).

9.2.17 a) $\{x_1, x_2\}$.

b) $\{x_1, x_2\}$.

c) $\{x_3\}$.

d) $\emptyset$.

e) $\mathrm{FV}(\varphi \wedge \psi) = \mathrm{FV}(\varphi) \cup \mathrm{FV}(\psi) = \{x_1\}$.

f) $\mathrm{FV}(\varphi \vee \psi) = \mathrm{FV}(\varphi) \cup \mathrm{FV}(\psi) = \{x_1\}$.

9.2.18 If $\varphi = (t_1 \doteq t_2)$ the result follows from Exercise 9.1.18. If $\varphi = \top$ or $\varphi = \bot$ the substitution does not change anything. In the cases $\varphi = (\varphi_1 \wedge \varphi_2)$, $\varphi = (\varphi_1 \vee \varphi_2)$ and $\varphi = (\varphi_1 \to \varphi_2)$ the result follows immediately from the inductive hypothesis, since substitution in such expressions are done by substituting in every place, and the inductive hypothesis says that the result holds in such cases. If $\varphi = \forall x_i \psi$ and x_j does not occur freely in φ, then $i = j$ or x_i does not occur freely in ψ. In the first case, this is clear since substitution does not change anything. In the second case, the result follows from the inductive hypothesis. Finally, in the case of $\exists$ one does the same as with $\forall$.

9.2.19 $\mathrm{FV}(\varphi) = \{x_1, x_3\}$.

$\varphi[f_1(x_3)/x_1] =$
$\forall x_2(\forall x_1 P_1(x_1, x_2) \to \exists x_2(f_1(f_1(x_3)) \doteq f_2(x_2, x_3))) \vee \forall x_3 \neg(f_1(x_3) \doteq x_3)$.

$\varphi[x_1/x_2] = \varphi$.

$\varphi[f_2(x_1, x_3)/x_3] =$
$\forall x_2(\forall x_1 P_1(x_1, x_2) \to \exists x_2(f_1(x_1) \doteq f_2(x_2, f_2(x_1, x_3)))) \vee \forall x_3 \neg(x_1 \doteq x_3)$.

9.2.20 a) $f_1 \doteq f_1$, $P_1(f_1(f_2))$, $P_1(f_2)$.

b) $x_0 \doteq x_1$, $x_1 \doteq x_0$, $P_1(f_1(x_0)) \doteq x_1$.

9.2.21 We start by showing that $t[x_i/x_i] = t$ holds for all terms t and variables x_i. If $t = x_j$ there are two cases to check. When $i = j$ we get $x_j[x_i/x_i] = x_i$, and when $i \neq j$ we get $x_j[x_i/x_i] = x_j$. In both cases, the result is equal to t. If $t = f_j(\ldots)$, the result follows immediately from the induction step.

We now show that $\varphi[x_i/x_i] = \varphi$ is true for all formulas φ and variables x_i. For atomic formulas, it follows from the fact that the corresponding property holds for terms, as we just proved. For composite formulas, it follows immediately from the inductive hypothesis, except for the case in which φ is of the form $\forall x_j \psi$ or $\exists x_j \psi$. We consider the first of these cases, as the other one is completely analogous. When $i = j$ we get, by definition of substitution, that $\varphi[x_i/x_i] = \varphi$. When $i \neq j$, we get $\varphi[x_i/x_i] = \forall x_j(\psi[x_i/x_i])$. Since the inductive hypothesis gives $\psi[x_i/x_i] = \psi$, the result follows.

9.2.22 Induction again. For atomic formulas, this reduces to Exercise 9.1.19. If $\varphi = \top$ or $\varphi = \bot$ the result is obvious since substituting does not change anything. If $\varphi = (\varphi_1 \wedge \varphi_2)$ the result follows immediately from the inductive hypothesis, as well as in the case of $\vee$ and $\to$. Consider now the case $\varphi = \forall x_i \psi$. Assume that y does not occur in φ; then we have $y \neq x_i$. We consider two cases, depending on whether $x = x_i$ or not. In the case $x = x_i$ we have $\varphi[y/x][x/y] = (\forall x_i \psi)[x/y] = \varphi[x/y] = \varphi$, where the last step uses that y does not occur in φ. Consider now the case $x \neq x_i$. The previous exercise handled the case where $y = x$, so we assume now that we have $y \neq x$. We then have $\varphi[y/x][x/y] = (\forall x_i \psi[y/x])[x/y] = \forall x_i \psi[y/x][x/y] = \forall x_i \psi$, where the last step uses the inductive hypothesis. In the same way one handles the case $\exists$.

9.2.23 With $f = \forall x_0(x_1 \doteq x_1)$, $y = x_0$, $x = x_1$ we have $\varphi[y/x][x/y] = \forall x_0(x_1 \doteq x_1)[x_0/x_1][x_1/x_0] = \forall x_0(x_0 \doteq x_0)[x_1/x_0] = \forall x_0(x_0 \doteq x_0)$.

10.1.1 If P_j is nullary, then its interpretation will be a proposition which is either true or false. If f_j is nullary, its interpretation will be an element of the domain (a constant).

10.1.17 a) $\mathcal{A}[x_i \mapsto a][x_i \mapsto b] = \mathcal{A}[x_i \mapsto b]$

b) $\mathcal{A}[x_i \mapsto [\![x_i]\!]^{\mathcal{A}}] = \mathcal{A}$

c) $\mathcal{A}[x_i \mapsto [\![x_i]\!]^{\mathcal{A}[x_i \mapsto b]}] = \mathcal{A}[x_i \mapsto b]]$

10.1.18 Assume that $i \neq j$. It is sufficient to check that $v[x_i \mapsto a][x_j \mapsto b](x_k) = v[x_j \mapsto b][x_i \mapsto a](x_k)$ holds for all variables x_k. If $k = i$ we get $v[x_i \mapsto a][x_j \mapsto b](x_k) = v[x_i \mapsto a](x_i) = a$ and $v[x_j \mapsto b][x_i \mapsto a](x_k) = a$.

If $k = j$ we get $v[x_i \mapsto a][x_j \mapsto b](x_k) = b$ and $v[x_j \mapsto b][x_i \mapsto a](x_k) = v[x_j \mapsto b](x_j) = b$. For every other k we get $v[x_i \mapsto a][x_j \mapsto b](x_k) = v(x_k)$ and $v[x_j \mapsto b][x_i \mapsto a](x_k) = v(x_k)$.

If, on the other hand, $i = j$, the terms are simplified, according to previous exercise, to $\mathcal{A}[x_j \mapsto b]$, respectively $\mathcal{A}[x_i \mapsto a]$. Thus, x_i is given the value b, respectively a by these valuations, so if $a \neq b$ then the valuations are not equal.

10.1.21 $[\![\forall x_0(x_0 \doteq x_1)]\!] = 1$ if and only if the domain has precisely one individual.

$[\![\exists x_0(x_0 \doteq x_1)]\!] = 1$ always holds.

10.2.2 We will show that if $[\![\forall x\varphi]\!] = 1$, then $[\![\varphi]\!] = 1$. The former means, by definition, that $[\![\varphi]\!]^{[x \mapsto a]} = 1$ holds for every element a in the domain, in particular, for $a = [\![x]\!]$. But then $[\![\varphi]\!] = [\![\varphi]\!]^{[x \mapsto [\![x]\!]]} = 1$.

10.2.3 We will show that if $[\![\varphi]\!] = 1$, then $[\![\exists x\varphi]\!] = 1$, i.e., that $[\![\varphi]\!]^{[x \mapsto a]} = 1$ for some choice of a. But if we let $a = [\![x]\!]$, then that follows immediately, since $[\![\varphi]\!]^{[x \mapsto [\![x]\!]]} = [\![\varphi]\!] = 1$. (Note that this argument is completely dual to the solution of the previous exercise.)

10.2.7 With $\psi = \neg\varphi$ we get the following proposition to consider: $\forall x_0(\varphi \vee \neg\varphi) \vDash \forall x_0\varphi \vee \neg\varphi$. But $[\![\forall x_0(\varphi \vee \neg\varphi)]\!] = 1$, while $[\![\forall x_0\varphi \vee \neg\varphi]\!]$ does not have to be 1. If $\varphi = (x_0 \doteq x_1)$ and the domain consists of at least two elements, it becomes, for instance $[\![\forall x_0\varphi \vee \neg\varphi]\!] = \neg[\![\varphi]\!]$, which is 0 if x_0 and x_1 are given the same value.

10.2.8 By taking $\varphi = \bot$ we get the following propositions to consider: $\forall x_0\bot \vee \psi \vDash \forall x_0(\bot \vee \psi)$. We will investigate how the truth values differ, so that we can simplify the formula to some other with the same truth value: $\forall x_0\bot \vee \psi$ has the same truth value as ψ and $\forall x_0(\bot \vee \psi)$ has the same truth value as $\forall x_0\psi$. We therefore have the following proposition to consider: $\psi \vDash \forall x_0\psi$. But this has been studied in Example 10.2.6.

11.1.3 Assume that $[\![\exists x\neg\varphi]\!] = 1$. This means that $[\![\neg\varphi]\!]^{[x \mapsto a]} = 1$ for some a. Hence, for this a, $[\![\varphi]\!]^{[x \mapsto a]} = 0$, so that it does not hold $[\![\varphi]\!]^{[x \mapsto a]} = 1$ for every a, which means that $[\![\forall x\varphi]\!] = 0$, and hence $[\![\neg\forall x\varphi]\!] = 1$. We can show the other direction by following the argument backwards.

11.1.7 It has been previously shown that $\varphi \vDash \exists x\varphi$ (Example 11.2.30). We shall prove the converse. Assume therefore that $[\![\exists x\varphi]\!] = 1$; that is to say, $[\![\varphi]\!]^{[x \mapsto a]} = 1$ for some a. If x does not occur free in φ, then $[\![\varphi]\!] = [\![\varphi]\!]^{[x \mapsto a]} = 1$.

11.1.10 Assume that $[\![\exists x(\varphi \wedge \psi)]\!] = 1$, which means that $[\![\varphi \wedge \psi]\!]^{[x \mapsto a]} = 1$ for some a. Then, for this a, $[\![\varphi]\!]^{[x \mapsto a]} = 1$ and $[\![\psi]\!]^{[x \mapsto a]} = 1$. But if x does not occur freely in ψ, we have $[\![\psi]\!]^{[x \mapsto a]} = [\![\psi]\!]$. Furthermore, since $[\![\varphi]\!]^{[x \mapsto a]} = 1$, $[\![\exists x\varphi]\!] = 1$. It follows that $[\![\exists x\varphi \wedge \psi]\!] = [\![\exists x\varphi]\!] \wedge [\![\psi]\!] = 1 \wedge 1 = 1$. The other direction is proven by following the previous argument backwards.

11.1.13 Yes, the formula is a tautology. Take an arbitrary interpretation $\mathcal{A}$. We consider two cases. Assume first that $P_1^{\mathcal{A}}(a)$ is false for a certain a in the domain. Then $[\![P_1(x_0) \to \forall x_0 P_1(x_0)]\!]^{[x_0 \mapsto a]} = 1$, and thus $[\![\exists x_0(P_1(x_0) \to \forall x_0 P_1(x_0))]\!] = 1$. Assume, on the other hand, that $P_1^{\mathcal{A}}(a)$ is true for all a. Then $[\![\forall x_0 P_1(x_0)]\!] = 1$, so that $[\![P_1(x_0) \to \forall x_0 P_1(x_0)]\!] = 1$ and hence $[\![\exists x_0(P_1(x_0) \to \forall x_0 P_1(x_0))]\!] = 1$.

11.1.15 $\neg\forall x(\varphi \to \psi) \approx \exists x\neg(\varphi \to \psi) \approx \exists x\neg(\neg\varphi \vee \psi) \approx \exists x(\varphi \wedge \neg\psi)$.

$\neg\exists x(\varphi \wedge \psi) \approx \forall x\neg(\varphi \wedge \psi) \approx \forall x(\neg\varphi \vee \neg\psi) \approx \forall x(\varphi \to \neg\psi)$.

11.2.13 $(\exists x_1 P_1(x_0, x_1))[x_1/x_0] = (\exists x_1 P_1(x_1, x_1))$. Its truth value is 1 if and only if $P_1^{\mathcal{A}}(a, a) = 1$ for some a in the domain. If $\mathcal{A}$ is the interpretation that is used in the example, this means that $a < a$ should be true for some real number. But it is not, so (11.2.12) is false in that interpretation.

11.2.15 Let $\varphi = \exists x_1 P_1(x_0, x_1)$, $t = x_1$, $j = 0$. Then the left hand side of (11.2.14) is the truth value of (11.2.12), which is 0 according to the previous exercise.
The right hand side is, on the other hand,

$$[\![\exists x_1 P_1(x_0, x_1)]\!]^{[x_0 \mapsto [\![x_1]\!]]},$$

which is 1 since $[\![P_1(x_0, x_1)]\!]^{[x_0 \mapsto [\![x_1]\!]][x_1 \mapsto [\![x_1]\!]+1]} = 1$, that means the same as $[\![x_1]\!] < [\![x_1]\!] + 1$, which is in turn true.

11.2.20 x_1 is free for x_0 in $\exists x_0 P_1(x_0, x_1)$ since x_0 is bound by $\exists$. This means that the condition $i \neq j$ in Definition 11.2.18 is not satisfied.

x_0 is bound for x_1 in the same formulas, since the variable x_0 is bound by the quantifier if one inserts it in the place of x_1. More formally, it holds that $1 \neq 0$, that x_0 occurs freely in $P_1(x_0, x_1)$ and that x_0 occurs in x_0. Therefore, the conditions in Definition 11.2.18 are satisfied.

x_0 is free for x_1 in $\forall x_0 P_1(x_0)$, since the formula does not contain x_1.

11.2.22 The substitution of x for x does not change anything, so it is clear that if the notion *free for* is defined in the right way, x should be free for x in φ. Checking this is done by induction. This is clear for atommic formulas, and for formulas composed by connectives it follows immediately by the inductive hypothesis. For formulas formed with $\forall$ and $\exists$ it follows immediately from the inductive hypothesis and the fact that the conditions "$i \neq j$" and "x_j occurs in x_i" in the definition of *bound for* cannot be satisfied simultaneously.

11.2.23 This is proved by induction. It is vacuously true for atomic formulas, since t cannot be bound for x in those. For formulas composed by connectives it follows from the inductive hypothesis. For quantified formulas it follows from the condition saying that x_j should occur in t.

11.2.24 It follows immediately from the previous exercise: if t was bound for x in φ, then some of the variables in t would be quantified in φ.

11.2.31 Assume that $[\![\forall x \varphi]\!] = 1$; that is, $[\![\varphi]\!]^{[x \mapsto a]} = 1$ for all a. If t is free for x in φ, we have therefore $[\![\varphi[t/x]]\!] = [\![\varphi]\!]^{[x \mapsto [\![t]\!]]} = 1$.

11.2.32 With $x = x_0$, $t = x_1$, $\varphi = \forall x_1 (x_0 \doteq x_1)$ we get the following proposition to consider: $\forall x_1 (x_1 \doteq x_1) \vDash \exists x_0 \forall x_1 (x_0 \doteq x_1)$. The first formula is true in all interpretations, but the other formula is only true if there is precisely one individual.

11.2.33 With $x = x_0$, $t = x_1$, $\varphi = \exists x_1 \neg(x_0 \doteq x_1)$ we get the following proposition to consider: $\forall x_0 \exists x_1 \neg(x_0 \doteq x_1) \vDash \exists x_1 \neg(x_1 \doteq x_1)$. The first formula is true if there are at least two individuals; the other one, on the other hand, is always false.

11.2.37 $[\![t[y/x][x/y]]\!] = [\![t[y/x]]\!]^{[y \mapsto [\![x]\!]]} = [\![t]\!]^{[y \mapsto [\![x]\!]][x \mapsto [\![y]\!]^{[y \mapsto [\![x]\!]]}]}$. But since $[\![y]\!]^{[y \mapsto [\![x]\!]]} = [\![x]\!]$, so we can continue: $= [\![t]\!]^{[y \mapsto [\![x]\!]][x \mapsto [\![x]\!]]} = [\![t]\!]^{[y \mapsto [\![x]\!]]}$.

One computes similarly for φ, but in the first equality, the condition that x is free for y in $\varphi[y/x]$ is needed, and in the second equality one uses that y is free for x in φ. We get the answer $[\![\varphi]\!]^{[y \mapsto [\![x]\!]]}$.

12.1.6 Let $\varphi = (u \doteq x)$, where x is a variable which does not occur in u. Then

$\varphi[t/x] = (u \doteq t)$ and $\varphi[s/x] = (u \doteq s)$. We can, hence, use the replacement rule.

12.1.8 End the derivation by using $\to I$ and three $\forall I$. To derive $x_0 \doteq x_2$, the replacement rule is used, and above it $\wedge E$.

12.2.1 End with $\to I$. To derive $\bot$, we use that from $\forall x\varphi$ we can derive φ by $\forall E$, and later $\exists x\varphi$ by $\exists I$. This gives a contradiction.

12.2.2 End with $\wedge I$, and above it use $\to I$. For $\forall x\varphi \vdash \varphi$ only one instance of $\forall E$ is needed. For $\varphi \vdash \forall x\varphi$ use $\forall I$, which is possible because x does not occur freely in φ.

12.2.3 Conclude as in the previous exercise. For $\exists x\varphi \vdash \varphi$, one instance of $\exists E$ is used, which is allowed because x does not occur freely in φ. For the other direction we use $\exists I$.

12.2.4 End with $\wedge I$ and thereafter $\to I$. For $\forall x(\varphi \vee \psi) \vdash \forall x\varphi \vee \psi$ RAA is used as the last step. To get to $\bot$ we use $\to E$ with $\neg(\forall x\varphi \vee \psi)$ as the main premise. The side premise is derived from φ, by $\forall I$ (this step requires that x does not occur freely in any undischarged assumption, but this condition is satisfied, since x does not occur freely in ψ) followed by $\vee I$. Finally, the formula φ is derived from $\forall x(\varphi \vee \psi)$ and $\neg(\forall x\varphi \vee \psi)$.

For $\forall x\varphi \vee \psi \vdash \forall x(\varphi \vee \psi)$ one discharges $\forall I$ (which requires that x does not occur freely in ψ). Thereafter, an instance of $\vee E$.

12.2.5 For $\exists x(\varphi \wedge \psi) \vdash \exists x\varphi \wedge \psi$ one discharges by $\exists E$, which is possible since x does not occur freely in ψ. The side derivation ends with $\wedge I$.

To derive $\exists x(\varphi \wedge \psi)$ from $\exists x\varphi \wedge \psi$ one ends with $\exists E$ applied to $\exists x\varphi$, which is possible since x does not occurs freely in ψ. The formula $\exists x\varphi$ is derived, in turn, using $\wedge E$ from $\exists x\varphi \wedge \psi$. To derive $\exists x(\varphi \wedge \psi)$ from φ and $\exists x\varphi \wedge \psi$ one ends with $\exists I$ and above it $\wedge I$. The formula ψ is derived through $\wedge E$ from $\exists x\varphi \wedge \psi$.

13.1.17 If $\varphi \vdash \forall x\varphi$, then, according to the soundness theorem, we would have $\varphi \vDash \forall x\varphi$, but if one takes $\varphi = (x_0 \doteq x_1)$ and gives the same value to both x_0 and x_1, then $[\![\varphi]\!] = 1$, while $[\![\forall x\varphi]\!] = 0$ if there are at least two individuals.

13.1.19 Assume that this set was inconsistent. Then, according to the soundness theorem, it could not have a model. But the following is a model of it: let the domain consist of two elements and let all variables have the same value.

13.3.2 a) From $\forall x\varphi$ one derives $\varphi[y/x]$ in one step if y is free for x in φ. One then concludes through $\forall I$ and asserts that $\forall y\varphi[y/x]$, which is correct if y does not occur freely in φ.

b) Let $\varphi = \exists y\neg(x \doteq y)$, where x, y are different variables. If we had (13.3.3) in this case, then according to the soundness theorem we should have $\forall x\exists y\neg(x \doteq y) \vDash \forall y\exists y\neg(y \doteq y)$. But the left hand side is true if the domain has at least two elements, while the right hand side is always false.

c) Let $\varphi = (y \doteq z)$, where x, y, z are different variables. If we had (13.3.3) in this case, then according to the soundness theorem we should have $\forall x(y \doteq z) \vDash \forall y(y \doteq z)$. But if one lets y and z have the same value and there are at least two elements in the domain, the left hand side is true while the right hand side is false.

13.3.4 a) From $\varphi[y/x]$ we derive $\exists x\varphi$ in one step, as long as y is free for x in φ. Therefore we can, by $\exists E$, conclude that $\exists x\varphi$ and discharge the assumption $\varphi[y/x]$, assuming that y does not occur freely in φ.

b) Let $\varphi = \forall y(x \doteq y)$, where x, y are different variables. If (13.3.5) was true, then according to the soundness theorem we would have $\exists y \forall y(y \doteq y) \vDash \exists x \forall y(x \doteq y)$. But the left hand side is always true, while the right hand side is true only if there is precisely one element in the domain.

c) Let $\varphi = (y \doteq z)$, where x, y, z are different variables. If (13.3.5) was true, then according to the soundness theorem we would have $\exists y(y \doteq z) \vDash \exists x(y \doteq z)$. But the left hand side is a tautology, while the right hand side is false if y and z have different values.

13.3.11 That y is free for x in $\forall y\psi$ means that x does not occur freely in $\forall y\psi$. This, in turn, means that $x = y$ or that x does not occur freely in ψ. In the first case, the result follows from Exercise 9.2.21. In the second case, the result follows from Exercise 9.2.18.

13.3.12 The case $y = x$ is already considered in Exercise 9.2.21, so we assume therefore that $y \neq x$. We prove the claim by induction. For atomic formulas, it reduces to Exercise 9.1.19. If $\varphi = \top$ or $\varphi = \bot$, the claim is obvious, since substitution does not change anything. If $\varphi = (\varphi_1 \wedge \varphi_2)$, the result follows immediately from the inductive hypothesis, as in the cases of $\vee$ and $\rightarrow$. Consider now the case $\varphi = \forall x_i\psi$. That y is free for x in φ means that x does not occur freely in φ or that both of the following assertions are true: $x_i \neq y$ and y is free for x in ψ. That y does not occur freely in φ means that $y = x_i$ or that y does not occur freely in ψ. Together, these two assumptions lead us to the following possible four situations:

a) x does not occur freely in φ and $y = x_i$.

b) x does not occur freely in φ and y does not occur freely in ψ.

c) $x_i \neq y$, y is free for x in ψ and $y = x_i$.

d) $x_i \neq y$, y is free for x in ψ and y does not occur freely in ψ.

Consider first case a. Then $\varphi[y/x][x/y] = \varphi[x/y] = \varphi$.

Consider now case b. Then $\varphi[y/x][x/y] = \varphi[x/y]$. If $y = x_i$ the result follows immediately, since substitution does not change anything. If $y \neq x_i$, then $\varphi[x/y] = \forall x_i\psi[x/y] = \forall x_i\psi = \varphi$.

In case c we have $\varphi[y/x][x/y] = (\forall y\psi[y/x])[x/y] = \forall y\psi[y/x] = \varphi$, where the last step is justified by the previous exercise.

In case d we split into two cases. If $x = x_i$ we have $\varphi[y/x][x/y] = \forall x_i\psi[x/y] = \forall x_i\psi = \varphi$. If $x \neq x_i$, we have $\varphi[y/x][x/y] = (\forall x_i\psi[y/x])[x/y] = \forall x_i\psi[y/x][x/y] = \forall x_i\psi$, where the last step follows by the inductive hypothesis.

13.3.13 a. Under these assumptions ψ can be derived by going from $\forall x_0\varphi$ with $\forall E$ to $\varphi[x_1/x_0]$ and then using $\forall I$ till $\forall x_1\varphi[x_1/x_0]$, ending afterwards with $\rightarrow I$. Then the soundness theorem gives that ψ is a tautology.
b. Take $\varphi = (x_1 \doteq x_2)$.
c. Take $\varphi = \exists x_1 \neg(x_0 \doteq x_1)$.
d. Take $\varphi = \forall x_1(x_0 \doteq x_1) \wedge (x_1 \doteq x_1)$.

13.3.14 a) The application of $\exists I$ is not correct, since the formula in the row above must be of the form $(x_1 \doteq x_1)[t/x_1]$ for some term t, but then it is $t \doteq t$ which is incorrect, since it should be x_0 to the left and x_1 to the right.

b) Yes. Use "refl", followed in the next row by $\exists I$, and finally $\rightarrow I$ (which does not discharge any assumptions).

c) No. According to the soundness theorem, it would then be true in all interpretations, while it is false in $\langle \mathbb{N}; >; \rangle$.

14.1.6 It suffices to check that $\Gamma \cup \{\neg\varphi\}$ is consistent. But if it was inconsistent, we could end with RAA and then deduce φ from Γ, which is impossible since Γ is closed under derivation and $\varphi \notin \Gamma$.

14.1.10 a) $f_1(x_0, x_1) \doteq x_1$. It belongs to Γ since $[\![f_1(x_0, x_1) \doteq x_1]\!] = 1 \iff [\![f_1(x_0, x_1)]\!] = [\![x_1]\!] \iff 0 + 1 = 1$.

b) Γ is consistent by the soundness theorem since it has a model, namely $\langle \mathbb{Z}; ; +, 0\rangle$. To check that it is maximally consistent, it is enough to check that $\Gamma \cup \{\varphi\}$ is consistent, so $\varphi \in \Gamma$. Assume therefore that $\Gamma \cup \{\varphi\}$ is consistent. If $\varphi \notin \Gamma$, then φ would be false in the interpretation, and hence we would have $\neg\varphi \in \Gamma$, which is not possible (refer to the Exercise 8.1.13).

c) No; we have $\exists x_0(f_1(x_0, x_1) \doteq f_2) \in \Gamma$, but $(f_1(t, x_1) \doteq f_2) \notin \Gamma$ for every term t. To see that this is the case, note that $[\![f_1(x_0, x_1) \doteq f_2]\!]^{[x_0 \mapsto a]} = 1 \iff a+1 = 0$. With $a = -1$ the formula is true, and hence the existential formula is true in the model. But $f_1(t, x_1) \doteq f_2$ has truth value 1 only if $[\![t]\!]+1 = 0$, which is impossible since no term has negative value in the model. This can be checked through an inductive proof: for terms that are variables, we have $[\![x_i]\!] = i$, which is non negative. For terms of the form $f_1(t, s)$, we have $[\![f_1(t, s)]\!] = [\![t]\!] + [\![s]\!]$, which is non negative since both $[\![t]\!]$ and $[\![s]\!]$ are non negative, by inductive hypothesis. For terms of the form f_2 the value is 0.

14.2.2 Because of theorem 14.1.2 it is enough to show that the needed formulas are derivable from Γ^*, since then they must be in Γ^*. That $\sim$ is reflexive follows from the fact that $t \doteq t$ can be derived by the rule "refl". That $\sim$ is symmetric follows since $t \doteq s \vdash s \doteq t$ (Example 12.1.3) and that it is transitive follows from Exercise 12.1.6.

14.2.3 As indicated, it suffices to derive $f_i(t_1, \ldots, t_{a_i}) \doteq f_i(s_1, \ldots, s_{a_i})$ from $t_j \doteq s_j$, for $j = 1, \ldots, a_i$, since the latter formulas are by assumption in Γ^*. Use first the reflexivity rule to derive $f_i(t_1, \ldots, t_{a_i}) \doteq f_i(t_1, \ldots, t_{a_i})$. Use then the replacement rule, with $t_1 \doteq s_1$, to change the first argument on the right hand side of s_1. Continue then with the substitution rule, a total number of a_i times, until all the arguments have been changed. The result is a derivation, all of whose unfinished assumptions are in Γ^*, and being Γ^* closed under derivations, it contains the final formula as well.

14.2.4 We will show by induction that $[\![t]\!] = \tilde{t}$. For variables it follows by definition, since $v(x_i)$ is defined as the equivalence class of x_i. Let us now carry on the induction step. Assume that $t = f_i(t_1, \ldots, t_{a_i})$. We have $[\![t]\!] = f_i^{\mathcal{A}}([\![t_1]\!], \ldots, [\![t_{a_i}]\!])$. According to the inductive hypothesis, the arguments are equal to $\tilde{t_1}, \ldots, \tilde{t_{a_i}}$, so the definition of $f_i^{\mathcal{A}}$ gives that $[\![t]\!]$ is the equivalence class of $f_i(t_1, \ldots, t_{a_i})$.

14.2.5 $t \in [\![t]\!]$ is, according to the previous exercise, equivalent to $t \sim t$, which is true since $\sim$ is refexive. (Exercise 14.2.2).

14.2.9 $[\![\exists x\varphi]\!] = 1$ is equivalent to $[\![\forall x\neg\varphi]\!] = 0$, which according to the previous lemma is equivalent to $[\![\neg\varphi[t/x]]\!] = 0$ for any term t which is free for x in φ. But $[\![\neg\varphi[t/x]]\!] = 0$ is equivalent to $[\![\varphi[t/x]]\!] = 1$.

14.2.15 Assume that Γ is consistent; we must show that it has a model. Suppose that it does not have any model. This would make $\Gamma \vDash \bot$ hold, since every model of Γ would be a model of $\bot$ (since there are no such models). But because of the completeness theorem, we would then have $\Gamma \vdash \bot$, and so Γ would be inconsistent, a contradiction.

Index